UNCERTAINTY ANALYSIS IN UPSCALING WELL LOG DATA BY MARKOV CHAIN MONTE CARLO METHOD

A Thesis

by

KYUBUM HWANG

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

May 2009

Major Subject: Geophysics

UNCERTAINTY ANALYSIS IN UPSCALING WELL LOG DATA BY MARKOV CHAIN MONTE CARLO METHOD

A Thesis

by

KYUBUM HWANG

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Approved by:

Chair of Committee,	Richard L. Gibson
Committee Members,	Yuefeng Sun
	Thomas A. Blasingame
Head of Department,	Andreas K. Kronenberg

May 2009

Major Subject: Geophysics

ABSTRACT

Uncertainty Analysis in Upscaling Well Log data By Markov Chain
Monte Carlo Method. (May 2009)
Kyubum Hwang, B.S., Yonsei University
Chair of Advisory Committee: Dr. Richard L. Gibson

More difficulties are now expected in exploring economically valuable reservoirs because most reservoirs have been already developed since beginning seismic exploration of the subsurface. In order to efficiently analyze heterogeneous fine-scale properties in subsurface layers, one ongoing challenge is accurately upscaling fine-scale (high frequency) logging measurements to coarse-scale data, such as surface seismic images. In addition, numerically efficient modeling cannot use models defined on the scale of log data. At this point, we need an upscaling method replaces the small scale data with simple large scale models. However, numerous unavoidable uncertainties still exist in the upscaling process, and these problems have been an important emphasis in geophysics for years. Regarding upscaling problems, there are predictable or unpredictable uncertainties in upscaling processes; such as, an averaging method, an upscaling algorithm, analysis of results, and so forth.

To minimize the uncertainties, a Bayesian framework could be a useful tool for providing the posterior information to give a better estimate for a chosen model with a conditional probability. In addition, the likelihood of a Bayesian framework plays an important role in quantifying misfits between the measured data and the calculated parameters. Therefore, Bayesian methodology can provide a good solution for quantification of uncertainties in upscaling.

When analyzing many uncertainties in porosities, wave velocities, densities, and thicknesses of rocks through upscaling well log data, the Markov Chain Monte Carlo (MCMC) method is a potentially beneficial tool that uses randomly generated parameters with a Bayesian framework producing the posterior information. In addition, the method provides reliable model parameters to estimate economic values of hydrocarbon reservoirs, even though log data include numerous unknown factors due to geological heterogeneity. In this thesis, fine layered well log data from the North Sea were selected with a depth range of 1600m to 1740m for upscaling using an MCMC

implementation. The results allow us to automatically identify important depths where interfaces should be located, along with quantitative estimates of uncertainty in data. Specifically, interfaces in the example are required near depths of 1,650m, 1,695m, 1,710m, and 1,725m. Therefore, the number and location of blocked layers can be effectively quantified in spite of uncertainties in upscaling log data.

To my parents who nurtured me in faith, love, and perseverance

ACKNOWLEDGMENTS

Praise and glory be to the Almighty God and respect and gratitude be to Jesus. I would like to express my deepest appreciation to my advisor Dr.Gibson for his generous guidance and for giving me the opportunity to learn computational skills and professional study such as theoretical seismology and borehole acoustics. He has inspired and taught me with much patience over the course of master degree. I am forever grateful for his support and encouragement. His never-ending help and wise criticism has had a tremendous influence on my research.

I extend my sincere appreciation to each of my committee members, Dr.Yuefeng Sun and Dr.Thomas A. Blasingame for their valuable comments that led to an improvement of this thesis. Dr.Sun is a excellent professor who made me easily understand petroleum geology, and Dr.Blasingame showed me cordial interest on my research.

Above all, I feel thankful for all affection and encouragement that I received from my and my wife's family. They always gave me absolute trust and prayed for me. Particularly, my beloved mother and father's constant love has inspired me to accomplish this goal.

I also wish to thank the faculty, staff, and colleagues of the Geohysics Department for the contribution to my work. Special thanks are given to all my good friends in the lab; Hoa, Ravi, Rituparna, Jatara, and Chad. In particular, Hoa has been a kind supporter to let me learn a Mathematica program with giving me a good idea while working on research. A word of appreciation is also given to my company, Korea Gas corporation for giving me the most meaningful chance to study abroad in my life.

Above all, I would like to thank my two lovely daughters (Sunha and Yooha) who bring inexpressible happiness to me. They are really God's precious presents. I also wish to give mother-in-law thankfulness for taking care of my children when my second daughter was born. And last but not least, there is my wife, Jae-Eun who made me realize that there is something more important in my life than work and study. Thanks to her who has been together with me during the best and worst of times, I could stand up again with valor. Without her support and love, I could not have gotten through everything. During my degree, she continuously encouraged me to be able to get over all adversities even though she tremendously suffered from raising two daughters including a new born baby.

TABLE OF CONTENTS

Page

CHAPTER Page

LIST OF TABLES

LIST OF FIGURES

FIGURE Page

FIGURE Page

FIGURE Page

FIGURE Page

CHAPTER I

INTRODUCTION

1.1 Motivation and Objectives

The motivation of this thesis originated from two areas which are the technical approach to geophysical inversion and the quantification of uncertainties in the inversion. The former is with regard to upscaling well log data, and the latter is relevant to quantification of uncertainties inevitably existing in upscaling. During the past several decades, research in the domain of geophysics has advanced and widened, especially in the fields of seismic interpretation and data processing. These research advancements were achievable by technical developments with the advent of super high-speed and mass-storage computers. In spite of these improvements, however, there is still considerable uncertainty in data measurements, processing, interpretation, etc. due to the limitation of geological and technical analyses. Subsequently, recent geophysics emphasize reducing the uncertainties in geophysical regions from the data acquisition to the application. A goal of the upscaling of heterogeneous media is to simplify the earth model without changing the overall seismic wave field during wave propagation (Gold et al., 2000). By applying equivalent medium theory (Backus, 1962), layering on a finer scale may be replaced by creating a model with far thicker and fewer layers (Folstard and Schoenberg, 1992). As a way of blocking well log data, averaging method of individual properties in layers has been continuously developed by many geophysicists (Backus, 1962; Schoenberg and Muir, 1989; Folstard and Schoenberg, 1992; Mukerji et al., 2001; Tiwary, 2007). Although Mosegaard introduced a concept of Markov Chain Monte Carlo (MCMC) method in inverse problems (Mosegaard and Tarantola, 1995), the method was not applied to actual upscaling well log data. Therefore, there have been still many difficulties in confidently determining blocked layers. That is why my first objective is in upscaling

---

This thesis follows the style and format of Geophysics.

actual log data from the North Sea and making a better decision of layer boundaries to be upscaled.

Second, there have always been tremendous financial risk in exploration for petroleum reservoirs due to the remarkable costs in drilling. Because these high risks are financially difficult for companies, economic factors such as a rational appraisal of measured seismic data and a statistical interpretation of optimized parameters are absolutely getting more important. As a fundamental theory and a method, a Bayesian framework and an MCMC method were used to estimate the uncertainties and to obtain best-fit model parameters in the allowable data range. Recently, the technologies make it possible to get fast, precise, and reliable results from data sets containing uncertainties. Consequently, my second interest focuses on quantifying these uncertainty problems by using a computer simulation.

This thesis provides the methodology known as MCMC for upscaling log data and can be applied to other depth ranges which are abundant hydrocarbon areas. The suggested MCMC method would make a better contribution in obtaining reliable model parameters from quantification of uncertainties in upscaling.

1.2 Upscaling Problem of Well Log Data

Heterogeneities in the earth's subsurface result in considerable variations in the elastic constants in rocks. These inhomogeneous properties cause problems in seismic interpretation because of the complex relationship between wave amplitudes and heterogeneity. In addition, randomly varying rock components such as micro minerals, cracks, and saturation make it almost impossible to perfectly make a seismic interpretation and to solve inverse problems from the measured data. To get reliable results from data processing, therefore, the upscaling method relating the small scale to the large scale has been one of the most interesting emphases in geophysicists for a long time. In addition, the recent advances of fast and precise computing technology have increased the demand for quantitative predictions and higher resolutions in the earth model from effective upscaling methods.

1.2.1 Fundamentals

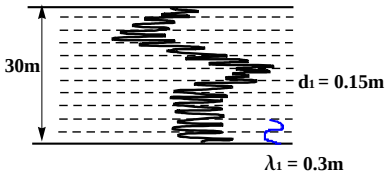
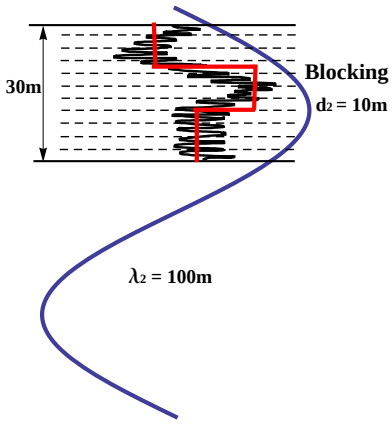
Upscaling means a mathematical prediction of the elastic properties in rocks

at lower frequency from higher frequency logging data on sonic velocities (Vp and Vs), porosity, and density (Bayuk et al., 2007). With respect to wave propagation, Stovas and Arntsen mentioned that the wave-propagation velocity strongly depends on the ratio (λ/d) of the dominant wavelength to the typical layer thickness (Stovas and Arntsen, 2006). When the wavelength is large enough comparing with the layer thickness, the wave velocity can be given by an average of the properties in individual layers (Backus, 1962), and waves behave as if propagating in an effective isotropic homogeneous medium even if the medium is anisotropic. In contrast, when the wavelength is small comparing with the layer thickness, the waves behave by the ray theory because of high frequency approximation. Namely, the application of ray tracing can be valid only for the media in which the spatial fluctuation in properties is larger than the wavelength (Cerveny et al., 1977).

Experimental research shows that sonic log data include a large fluctuation in P- wave and S- wave velocities, and this fluctuation decreases at the lower frequencies of the seismic data. Therefore, in a practical sense, Backus averaging means the replacement of a heterogeneous volume with a homogeneous volume containing effectively equivalent elastic constants. It is also a moving window process, whereby a stack of thin layers is averaged until it is approximated by the properties of a single thick layer. From this reason, it is reasonable to replace many thin layers with one thick layer.

With a lack of more easily detected reservoirs, recently found reservoirs have geologically more heterogeneous structures from micron size pores to several kilometers basins. To analyze these, we need to consider all heterogeneous effects on the small scale data, but it is very inefficient to resolve the every heterogeneity from numerous numerical simulations because of the excessive computing time. Therefore, the aim of geophysicist should be to make precise, timely, and cost saving decisions in all kinds of data analysis. By blocking well log data, we can considerably reduce the computation time and expenses when analyzing the given data. The upscaling method using long wavelength with respect to the layer thickness could be a reasonable method because the modeled heterogeneity of the small scale parameters will surprisingly decrease.

Table 1.1. Example of the comparison of well log data with upscaling log data. (In order to easily compare with each other, it is supposed that the same minimum velocity (V_p) and total layer thickness (T) were used.

	Well Log Data	Comparison	Upscaling Log Data
Frequency (f)	10 kHz	>>	30 Hz
Minimum (V_p)	3 km/s	=	3 km/s
Wavelength (λ)	$\lambda_1 = \frac{V}{f_1} = \frac{3,000}{10,000} = 0.3\,m$	<<	$\lambda_2 = \frac{V}{f_2} = \frac{3,000}{30} = 100\,m$
Total Layer Thickness (T)	30 m	=	30 m
Each Measured Thickness (d)	d_1 = 0.15m	<<	d_2 = 10m
Number of Layers (N)	$N_1 = \frac{T}{d_1} = \frac{30}{0.15} = 200$ Layers	>>	$N_2 = \frac{T}{d_2} = \frac{30}{10} = 3$ Layers
Figure		≠	
Ratio of λ to d	$\frac{\lambda_1}{d_1} = 2$	<	$\frac{\lambda_2}{d_2} = 10$
Computation Time	t_1	>>	t_2
Wave Behavior	Anisotropic	≠	Isotropic (Waves behave as if propagating isotropic homogeneous medium)

For example, as shown in Table 1.1, if the well log data are measured with a frequency of 10KHz, a P-wave velocity of $3km/s$, and a sample depth interval of 0.15m, the wavelength becomes 0.3m from the equation V/f. If we choose a depth range of 30m, the number of layers becomes 200, whereas the case of upscaling log data has only 3 layers and a large wavelength of 100m supposing that it uses low frequency (30Hz) waves and the same p-wave velocity ($3km/s$). Evidently, the latter is the more efficient method because it makes seismic computation time significantly reduced. However, a problem of quality in upscaling log data still remains because well log data show higher vertical resolution than one of seismic data. Therefore, seismic data need to be complemented with well log data, particularly with respect to important portions with low vertical resolution near reservoirs.

1.2.2 Previous Works

The ways of blocking well log data have been continuously developed as one part of an upscaling process. In finely layered media, the problem of elastic wave propagation had been the main interest of geophysicists, and Thomson(1950), Helbig(1958), and Anderson(1961) examined homogeneous anisotropic multilayered cases. Thomson gave a formal solution for waves of arbitrary wavelength in a medium with homogeneous isotropic layers, and he found the displacement and vertical stresses at any interface (Thomson, 1950). Helbig represented formulas for five elastic coefficients as averages and generalized them to a multilayered case, but he did not consider anisotropic layers (Helbig, 1958). After that, Anderson applied the formulas to anisotropic layered media (Anderson, 1961).

Despite these continuous efforts, we can say that a practical basis of the averaging method was made by Backus, because he showed that waves behave as if going through a homogeneous and transversely isotropic medium in case that the waves have a large enough wavelength comparing with a layer thickness (Backus, 1962). The Backus averaging approach is advantageous for two reasons. First, it gives simple expressions for elastic constants as averages of elastic moduli. Secondly, it can be applied to non-periodic layered media with more than two constituents, which may be transversely isotropic themselves. Also, it has been extended to the case where the constituents themselves are anisotropic (Schoenberg and Muir, 1989).

Assuming that the given medium consists of thin isotropic layers with P- and S-

wave velocity, and density, five elastic constants in the medium can be calculated by multiplying shear modulus and velocities (Schoenberg, 1983), and Hsu et al. (1988) used a moving average window to estimate seismic velocities from anisotropic sonic logs. In this moving average, a total depth is divided into the same thick zones, and each fine layer zone is replaced by its equivalent homogeneous anisotropic thick layer (Hsu et al., 1988). In a finely layered medium, sonic data indicate drift curves with measured depths due to effects of fine layering in each depth interval. Folstard mentioned that fine layering of one tenth of the smallest wavelength does not affect seismic wave propagation even if the medium is anisotropic (Folstard and Schoenberg, 1992).

Recently, Tiwary 2007) compared three upscaling methods; simple averaging, Backus averaging, and pair correlation function averaging. Simple averaging includes a static scale effect, and Backus averaging contains both a static scale effect and the interaction among layers. In addition, pair correlation function averaging includes a static scale, interaction, and scattering. He also mentioned that attenuation and dispersion cause multiple scattering at layer interfaces as well as heterogeneities. Although Backus did not consider this multiple scattering, Backus averaging can give a good solution in most cases using long wavelength because the fluctuation of P- and S- wave velocities in sonic data decreases at the lower frequencies (Tiwary, 2007). Prior to Backus averaging, however, we need to make two assumptions. One is that the media is horizontally stratified and azimuthal variations are ignored in this work. The other is that well log data with high resolution follow a TI (transversely isotropic) model which is isotropic at any vertical measuring points.

With these upscaling methods, we could understand how to solve upscaling problems numerically, but we still need to consider that exterior error factors could have an important effect on results of upscaling log data because fine-scale well log data inevitably contain error factors due to the limitations of technical measurements and geological analyses in heterogeneous media. An MCMC method is a computer-based algorithm of obtaining good models by randomly generating samples. The algorithm uses a way of analyzing the chosen models statistically with internally determined probabilities. Upscaling based on an MCMC method can freely control main factors including other input variables such as iteration and the number of blocked layers. In addition, results of upscaling by an MCMC method are considerably in good agreement to original well log data. Therefore, an MCMC method could give us more

predictable and reliable results of upscaling.

1.3 Uncertainty Problem in Upscaling

There are countless known or unknown uncertainties in upscaling well log data. In a geophysical view, how we quantify uncertainties is one of the most crucial problems in geophysicists who should analyze and evaluate well log data from field experiments to computational data processing.

1.3.1 Cause and Classification

In seismic analyses, uncertainties of data mainly come from remote measurements for unknown subsurface properties because irregular changes of layer thicknesses or rock porosities could cause uncertain variations in the reflection amplitudes. In addition, unremoved tuning effects may cause an anomalous amplitude in data processing. Furthermore, these uncertainties are not removed but kept or even enlarged during data processing such as upscaling log data. In order to quantitatively analyze uncertainties, a statistical approach such as histogram or probability density function (PDF) needs to relate subsurface geological properties to measured seismic data (Mukerji et al., 2001).

Since fine-scale measurements include large variations, a great deal of uncertainties could be found in seismic data calibration. Core samples provide information from micron to centimeter scale, and well log data cover from centimeter to meter scale, and seismic data indicate from meter to decimeter scale. In practice, damaged samples and washout in boring tool could be causes of inaccurate data.

Uncertainty in geophysics can be classified according to both geologic uncertainty and measurement uncertainty. Geologic uncertainty is very unpredictable because of the existence of numerous different rock properties. These properties relating to reflectors do not uniquely determine lithologies from reflectors. Therefore, it is very difficult to completely remove this uncertainty and to make an appropriate geological model for the interpretation (Houck, 2002).

Regarding upscaling problems, there are numerous predictable or unpredictable uncertainties in original fine-scale data, averaging methods, upscaling algorithms, computing processes, statistical analyses, etc. In case of well log data (Vp, Vs,

density, Caliper log, Gamma ray, etc.), finely layered media could basically include experimental or environmental uncertainties even if there are no human errors. For example, the uncertainties include non-vertical propagation of P- and S- waves, cracks of the rocks, and changes of saturation during the well logging. In addition, when blocking well log data, additional uncertainties could happen according to various averaging methods showing different results. Concerning upscaling algorithms and computation processes, they depend on an expert's skill or technical ways of implementing algorithms.

1.3.2 Quantification

In upscaling processes, coarse-scale models have unavoidable uncertainties because it is impossible to make a complete large scale model with exactly the same response as a finely layered model. All we can do in upscaling is only to minimize these uncertainties with statistical analyses and quantification. In that sense, a Bayesian framework using posterior probabilities and an MCMC algorithm analyzing misfit results could be good methodology to quantify uncertainties in upscaling log data. Since an MCMC algorithm selects the next model with a specific step size according to the prior information and the likelihood function, probabilities taking new models could be good indicators of the posterior probabilities. Therefore, an MCMC method could be a useful tool to quantify uncertainties in upscaling log data by means of statistical analyses of chosen models.

1.4 Inverse Problem

The goal of solving an inverse problem is to select optimized model parameters based on general knowledge and geophysical measurements. Conceptually, an inverse problem is a task where observed data should be changed into some model parameters by mathematical calculation. That is, it describes how a parameterized physical system can be derived from the observed data. Therefore, a solution of inversion produces model parameters from the given data, and the objective is to find an acceptable inference about model parameters based on seismic data and geological prior knowledge.

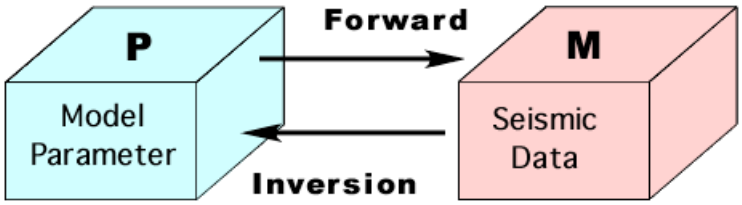


Fig. 1.1. The figure shows a conceptual diagram for a forward and an inversion problem.

As shown in figure 1.1, the key equation is

$$M = PV + e,$$

where M is a vector of the measurements m_i, P is a matrix of relating model parameters and data p_{ij}, V is a vector of the parameters v_j, and e is measurement errors. Inverse problem is to solve the equation for v_j which depends on the measurement system. For logging data, v_j might be a distribution of lithology and porosity.

Typically, the principle procedure of inverse problems is: first, model parameters describing geological bodies such as a reservoir should be set. Second, observed data, usually called as measurement data, should be obtained. Finally, the relationship between the observed data and the predicted parameters should be technically characterized.

Seismic data are often inverted for impedance consisting of velocity and density in order to estimate reservoir properties. Moreover, geophysical inverse problems often extend to quantifying uncertainties in model parameters. Unknown noise and uncertainties occurring in seismic measurements have a strong effect on results of inverse problems. Statistically, a probability density function (PDF) on a model space represents interpretable characteristics with solutions of inverse problems. A Bayesian theory is mainly chosen to analyze geophysical inverse problems, where it is possible to combine the prior information with the newly obtained information from observed data. (Mosegaard, 1998; Mosegaard and Tarantola, 1995).

To statistically estimate rock properties containing many uncertainties, an MCMC algorithm provides a good solution of inverse problems because an inversion code using an MCMC algorithm can generate stochastic models with an optimized posterior distribution.

1.5 Contribution

Major contributions of this thesis are the extension of an MCMC method to upscaling problems and advancement of quantification techniques. In practice, there are various uncertainties in upscaling log data originating from heterogeneous geological properties and technical limitations of measurement in a field. The core of an MCMC algorithm is that we can arbitrarily control the variables such as a σ value which means a 'noise' component and the number of layers which are upscaled. In addition, an MCMC method makes it possible to generate many reasonable models and to analyze them statistically. PDFs and histograms of the model parameters become good indicators to quantify uncertainties in upscaling log data.

With respect to the application of upscaling by an MCMC method, various depth ranges containing resourceful hydrocarbon areas could be investigated and evaluated in view of economic worth as well. Regarding methodological efficiency, an MCMC method could save time and expenses in analyzing uncertainties in model parameters by using a computer-based generating system of valid models. Furthermore, this method could facilitate the regulation of quantity and quality of new models in geophysical inverse problems.

1.6 Scope of this Thesis

The major content of this thesis consists of four chapters, and an upscaling method by an MCMC algorithm is introduced in detail as a main topic. Briefly, this thesis is organized as the following. In the first chapter, motivation and objectives are introduced briefly, and basic concepts of upscaling, uncertainty, and inverse problems are explained with main background knowledge and previous studies. The second chapter explains a concept and methodology of an MCMC method. In particular, the concept of an acceptance probability is mentioned more specifically to understand a main algorithm. The third chapter presents an actual upscaling procedure using an MCMC method which contains the determination of reasonable model parameters and the statistical analysis of upscaling data according to a computational algorithm. In addition, the chapter explains influential factors such as iteration, the number of layers, a σ value, and a coefficient of step size. To confirm accuracy of upscaling, the MCMC method was applied to other depth layers containing abundant hydrocarbon

areas.

Two appendices present the basic concept of a Bayesian framework and a simple example of an MCMC method. The Bayesian theory is introduced at length with the prior information, the likelihood function, and the posterior information, respectively. Furthermore, a Monte Carlo method is classified by a simple MC and an MCMC method, and the two methods were compared with each other by applying them to a simple polynomial problem. In order to facilitate the construction of a code, a main MCMC algorithm is summarized step-by-step as well.

CHAPTER II

MARKOV CHAIN MONTE CARLO (MCMC) METHOD

A simple Monte Carlo (MC) method is an inefficient algorithm of generating random samples, because it randomly tests a large number of unrelated models by comparing produced values to data. Each test model is chosen by randomly selected model parameters (see Appendix B). As an upgraded method, an MCMC method provides a more developed tool to find out a reasonable sample because the next sample is always generated depending on the most recent sample. It is more efficient, since it tests a large number of models that are likely to be good. In other words, this method provides an improved algorithm of generating random samples within a specific step size whenever iterating the loop of selecting samples. Furthermore, an MCMC method could be a good estimator of a reliable sample by quantifying chosen samples.

2.1 Concept

An MCMC method is a kind of algorithm for sampling from probability distribution by constructing a Markov chain which is a sequence of random variables $X_1, X_2, X_3, \cdots$ with Markov property. In a Markov chain, a newly chosen sample is only a function of the most recent value. That is, the future and past states are independent of present state.

$$P(X_{n+1} = x | X_n = x_n, \cdots, X_1 = x_1) = P(X_{n+1} = x | X_n = x_n)$$

The transition probability between two continuous samples depends on only the previous sample. More formally writing,

$$P(X_{n+1} = x | X_n = y) = P(X_n = x | X_{n-1} = y),$$

where $P(X_{n+1}|X_n)$ is a transition probability. An MCMC method gives us an efficient sampling algorithm based on a Markov chain.

MCMC has generally four important characteristics which are stationary distri-

bution, irreducibility, aperiodicity, and ergodicity. As time goes to infinity, a Markov chain converges to its invariant stationary distribution, and irreducibility means that any set of states can be reached from any other state. Assuming a stationary distribution exists, it is unique if the chain is irreducible. In addition, a chain is said to be aperiodic when the number of steps requiring to move between two states is not multiple of some integer. Moreover, a Markov chain is ergodic, namely, pertaining to the condition that a system will return to states which are closely similar to previous ones in an interval of sufficient duration. Because of this ergodic property, in theory, there is no need to run multiple chain with the property independent of initial or intermediate value (Sahu, 2000).

2.2 Methodology

An MCMC method is a sampling tool based on a Bayesian framework which is composed of the prior information, the likelihood, and the posterior information. If a model parameter vector and an observed data vector is **m** and **d**, respectively, we can represent the posterior probability as following the Bayesian rule.

$$P(m|d) = \frac{P(d|m)P(m)}{P(d)} = \frac{likelihood \times prior}{evidence}, \tag{2.1}$$

where $P(m|d)$ is known as 'the posterior', and $P(d|m)$ is referred to 'the likelihood'. If the evidence term has a unit probability by normalization, the likelihood is obtained by multiplying the observed information by the prior information.

When the observed data (d_{obs}) have Gaussian experimental uncertainties, the likelihood function can be written as following.

$$P(d|m) = \frac{1}{(2\pi)^{n/2}|C_d|^{1/2}} \ \ exp[-\frac{1}{2}(g(m) - d_{obs})^T C_d^{-1}(g(m) - d_{obs})], \tag{2.2}$$

where *g(m)* is a calculated value from model parameters, and C_d is a data covariance matrix. If correlated errors are not included in the data, the covariance matrix C_d can be estimated as $C_d = \sigma^2\mathbf{I}$ in which **I** is an identity matrix, and the σ should be estimated as a standard deviation of the errors. If, of course, error correlations exist without being neglected, the covariance matrix C_d must be applied in the inverse problem. The parameter set **m** means $\{m_1, m_2, m_3, \cdots\}$ in the modal space, and the measured data set represent as $\{d_{obs}^1, d_{obs}^2, d_{obs}^3, \cdots\}$ in the data space (Appendix A).

As a first step of MCMC method, we need an initial condition to define model parameter, observed data, and input variables as following.

$$\begin{cases} g(m) & : \text{ Parametric Model Function} \\ d_{obs} & : \text{ Observed Data} \\ n & : \text{ Iteration} \\ \sigma \text{ value} & : \text{ Control Variable,} \end{cases}$$

where the σ value is a very influential factor to control the likelihood and probabilities taking new models; it is defined mathematically below. Because this is a typical example of an inverse problem, other initial variables could be added or changed depending on characteristics of a problem. As the prior information, model parameters (m_{ij}) are selected with an arbitrary probability distribution function within an allowable parameter range. In addition, a way of choosing the next sample ($m_{i(j+1)}$) is adding α (known as step size in MCMC) to a previous sample (m_{ij}) as below.

$$m_{i(j+1)} = m_{ij} + \alpha \quad \begin{cases} i = [1, 2, 3, ..., p], & p: \text{ the number of parameters} \\ j = [1, 2, 3, ..., n], & n: \text{ iteration,} \end{cases}$$

where α is a uniformly distributed random number within an arbitrarily chosen step size. To decide which model is better or worse, we need the likelihood function ($L(m_{ij})$) including an error value ($S(m_{ij})$). The likelihood function is different depending on uncertainties of the observed data (d_{obs}). Supposing that the data have Gaussian experimental uncertainties, the likelihood function becomes the equation 2.3. If the experimental uncertainties follow a Laplacian function or a double Gaussian, the likelihood function should be changed to meet each function.

$$S(m_{ij}) = [\sum_{i=1}^{N} (g^i(m) - d^i_{obs})^2]^{\frac{1}{2}}$$

$$L(m_{ij}) = k \; exp[-\frac{1}{2}\frac{S^2(m_{ij})}{\sigma^2}], \tag{2.3}$$

where N is the number of data. With the likelihood function, we can determine a better sample as following.

$$m_{i(j+1)} = \begin{cases} \text{Better Model} & \text{if } L(m_{i(j+1)}) \geq L(m_{ij}) \\ \text{Worse Model} & \text{if } L(m_{i(j+1)}) < L(m_{ij}) \end{cases}$$

If the likelihood of a new sample is greater than the one of an old sample, the new sample is chosen as a better one. However, if a new sample is worse one, we need another step to decide which one we choose. We should sometimes select a new sample within the given acceptance probability (P_{accept}) even if the likelihood of a new sample is less than the one of an old sample, using the following probability models.

$$P_{accept} = \begin{cases} 1 & \text{if} \quad L(m_{i(j+1)}) \geq L(m_{ij}) \\ \exp(\text{-}\frac{\triangle S^2}{2\sigma^2}) & \text{if} \quad L(m_{i(j+1)}) < L(m_{ij}), \end{cases} \tag{2.4}$$

where $\triangle S^2 = S^2(m_{i(j+1)}) - S^2(m_{ij})$. In equation 2.3 and 2.4, we should notice that an error value ($S(m_{ij})$) and a σ value are very important factors to control the likelihood and the acceptance probability. In particular, we need to be careful to choose a σ value as an input variable because inaccurately chosen parameters can seriously skew a posterior probability density (PPD).

2.3 The Acceptance Probability

The probability of taking a new model is the most crucial factor determining how many and how precisely models are chosen. An MCMC algorithm largely consists of two portions. One is concerning effective generation of random parameters. The other is about the acceptance probability which plays an important role in deciding whether a new model will be accepted or rejected.

In Fig. 2.1, the acceptance probability is an exponential function of the ratio ($\frac{\triangle S^2}{2\sigma^2}$). Because the ratio value depends on two variables consisting of a square RMS error difference ($\triangle S^2 = S^2_{new} - S^2_{old}$) and a control variable (σ^2), trading off the two variables controls the acceptance probability. If $\triangle S^2$ keeps increasing, we need to catch it conceptually. Namely, it means that a square RMS error value of a new model gets larger than one of an old model with increasing iteration. In this case, the

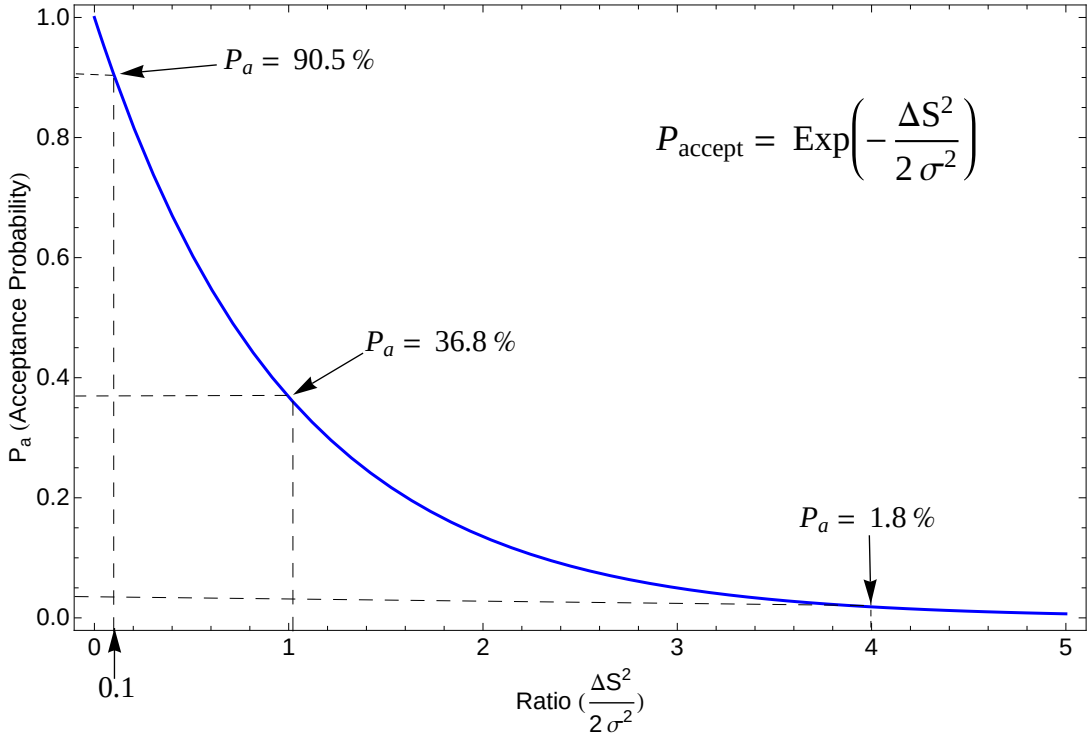


Fig. 2.1. The figure shows that the acceptance probability (P_a) strongly depends on the ratio of $\frac{\triangle S^2}{2\sigma^2}$ in the MCMC algorithm.

acceptance probability has a very small value, which also means that the probability taking a new model is very low. In addition, another important variable (σ^2) needs to be carefully treated because of a significant influence of the ratio ($\frac{\triangle S^2}{2\sigma^2}$) on the acceptance probability. From the given algorithm, the variable (σ^2) is not calculated but given as a kind of input condition. That is to say, it is a control factor to freely adjust the acceptance probability. If σ^2 increases larger than $\triangle S^2$, the ratio ($\frac{\triangle S^2}{2\sigma^2}$) will decrease and the probability taking a new model will increase. As a result, whether a new model is accepted or rejected strongly depends on not each change of $\triangle S^2$ and σ^2 but change of the ratio ($\frac{\triangle S^2}{2\sigma^2}$).

In Fig. 2.1, if the ratio ($\frac{\triangle S^2}{2\sigma^2}$) has value 1, the acceptance probability is 36.8 percent, but if the ratio decreases to one tenth, the probability increases to 90.5 percent. Whereas, if the ratio increase to four times, the probability decreases to 1.8 percent. Supposing that values of $\triangle S^2$ vary from 0 to 200, the probability of choosing a new model could be either over 90 percent or below 1 percent depending on the control factor (σ^2). Therefore, an MCMC algorithm can adjust at what rate

it selects new models with variations of square RMS error values.

2.4 Main Algorithm for Upscaling

The main MCMC algorithm we propose consists of setting input variables, building seismograms from chosen models, establishing a criterion to get better models, and generating posterior models from the acceptance probability. Ultimately, it will be followed by quantifying uncertainties in chosen models. As input variables, a σ value, iteration, the number of layers, and a coefficient of step size could be considered. A reference synthetic seismogram is obtained from well log data by computing the response using all samples in the well logs for velocity and density. Error is computed for a test. A calculated seismogram comes from randomly generated models with a small number of blocked (upscaled) layers. In addition, the rule producing new models shows that the next models are generated only within a specific step size, and a better model is could be determined by comparing with two likelihoods of an old and a new model. Finally, the acceptance probability depends on the likelihood, the difference between two square RMS errors ($\triangle S^2$), and the σ^2 value. By an MCMC method, reliable results based on the posterior distribution could be yielded, and uncertainties in upscaling data could be quantified with statistical analyses such as histograms and PDFs of accepted models.

Step 1. Input Variables

$$\begin{cases} \sigma : \text{Main Control Variable to determine acceptance probabilities} \\ \text{n} : \text{Iteration} \\ l_N\text{: The Number of Layers to be blocked} \\ \omega : \text{A Coefficient of step size,} \end{cases}$$

where we should note that σ is very correlated with RMS error value in a likelyhood ($L(\chi_{ij})$) and an acceptance probability (P_{accept}), and ω is used to adjust a step size.

Step 2. Calculate Seismogram from Each Model

$$\begin{cases} U^*(t) = f(V_p^*, V_s^*, \rho^*, \text{ thickness}^*, \text{ minimum offset, maximum offset, offset interval,} \\ \qquad \text{target depth, source frequency)} \\ \chi_{ij} = [\chi_{1j}, \chi_{2j}, \chi_{3j}, ..., \chi_{(l_N+1)j}] \\ U_j(t) = f(V_{pj}, V_{sj}, \rho_j, \text{ thickness}_j, \text{ minimum offset, maximum offset, offset interval,} \\ \qquad \text{target depth, source frequency),} \end{cases}$$

where * means the original data set, and χ_i is a boundary level of each layer. j means each step of iteration (j=1,2,3,...,n). U^* and U are used to compute error values in the below equation 2.7.

Step 3. The Prior information

The depth of the interfaces between the l_N layers are selected with an arbitrary probability distribution function between upper and lower layers.

Step 4. Rule to produce the next model

$$\chi_{i(j+1)} = \chi_{ij} + \alpha_{ij} \quad \begin{cases} i = [1, 2, 3, ..., l_N + 1] \\ j = [1, 2, 3, ..., n], \end{cases} \tag{2.5}$$

where α_{ij} is a uniformly distributed random number between $-\frac{\chi_{(i-1)j}+\chi_{(i+1)j}}{\omega}$ and $\frac{\chi_{(i-1)j}+\chi_{(i+1)j}}{\omega}$, and $\chi_{i(j+1)}$ is the next model of χ_{ij} which means interface depths.

Step 5. Criterion to decide which model is better or worse

$$L(\chi_{ij}) = k \; exp[-\frac{1}{2}\frac{S^2(\chi_{ij})}{\sigma^2}] \tag{2.6}$$

$$S(\chi_{ij}) = \sqrt{\frac{\sum_{i=1}^{N}\left[U_j(t)_i - U^*(t)_i\right]^2}{N}} \tag{2.7}$$

$$\chi_{i(j+1)} = \begin{cases} \text{Better Model} & \text{if } L(\chi_{i(j+1)}) \geq L(\chi_{ij}) \\ \text{Worse Model} & \text{if } L(\chi_{i(j+1)}) < L(\chi_{ij}) \ , \end{cases}$$

where σ^2 means data error variance, $L(\chi_{ij})$ is the likelihood of the model parameter χ_{ij}, $S(\chi_{ij})$ is RMS error used in the reference seismogram ($U^*(t)$) and the calculated seismogram ($U_j(t)$), and N is the number of calculated points in a seismogram.

Step 6 Generate the posterior models with comparison of the calculated seismogram with the original one using the Metropolis rule.

For the acceptance probability (P_{accept}),

$$P_{accept} = \begin{cases} 1 & \text{if} \quad L(\chi_{i(j+1)}) \geq L(\chi_{ij}) \ ----(1) \\ \exp(-\frac{\triangle S^2}{2\sigma^2}) & \text{if} \quad L(\chi_{i(j+1)}) < L(\chi_{ij}) \ ----(2), \end{cases}$$

where $\triangle S^2 = S^2(\chi_{i(j+1)}) - S^2(\chi_{ij})$.

$$\chi_{i(j+1)} = \begin{cases} \chi_{1j}, \chi_{2j}, \chi_{3j}, ..., \chi_{(l_N+1)j} & \text{in case of (1)} \\ \chi_{1(j+1)}, \chi_{2(j+1)}, \chi_{3(j+1)}, ..., \chi_{(l_N+1)(j+1)} & \text{in case of (2),} \end{cases}$$

where j=[1,2,3,...,n], and n=iteration.

$L(\chi_{ij})$ and $L(\chi_{i(j+1)})$ are calculated from χ_{ij} with $[\chi_{1j}, \chi_{2j}, \chi_{3j}, ..., \chi_{(l_N+1)j}]$ and $\chi_{i(j+1)}$ with $[\chi_{1(j+1)}, \chi_{2(j+1)}, \chi_{3(j+1)}, ..., \chi_{(l_N+1)(j+1)}]$, respectively.

As a first step, we determine input variables fit to our objective. These variables consist of a control variable (σ), the number of layers to be blocked (l_N), iteration, and a coefficient of step size (ω), where σ is a very importantly used variable to control probabilities taking new models. To get more new models, we can just increase a value of σ, but it is difficult to expect reliable models with a high value of σ. In a theoretical concept, σ^2 could be considered as 'noise' variance in seismic data. A higher value of σ will bring about ambiguous models with permitting larger error

values. Whereas, we can obtain well fitting models to upscaling with a low value of σ, but ultimately it also does not give confident results because the number of models considerably decrease and we can not estimate a quantified model. Therefore, reasonable control of σ needs to produce adequate and reliable models. With respect to iteration, generally larger iteration can produce more models except the case that results swiftly converge into a specific value and keep the same value because of a very small value of σ. Finally, a coefficient of step size (ω) could be an influence factor on selecting models as well. A step size here means an allowable size to move from a previous model to the next model. In this upscaling problem, an arbitrary interface depth (χ_i) could be generated between a top level of a upper layer and a bottom level of a lower layer with a specific step size. A larger value of ω means smaller range of walking (see step 4). However, this coefficient also should be carefully used because a very large value of ω can even prevent normal random walks and yield only a limited range of results.

In the second step, a randomly generated model consisting of interface depths is ultimately used to make a seismogram from Backus averaging of parameters (V_p, V_s, ρ, and thickness) within blocked layer levels in each iteration, and the calculated seismogram is used in step 5 to compare with the reference seismogram based on original well log data. The prior information in step 3 shows that arbitrary interface depths are randomly generated between upper and lower layers, and these indicate Gaussian distribution in PDF of each parameter. In step 4, the rule to generate the next models explains that a new model has a close relationship with a prior model as if two continuous models are connected with a specific step size. Step 5 provides a criterion of decision to choose a better model from comparison of two likelihoods. However, all worse models are not rejected, but some worse ones are accepted within an allowable probability as described in step 6.

CHAPTER III

UPSCALING BY AN MCMC METHOD

3.1 Background for Well Log Data

Most countries rely heavily on hydrocarbon products, and oil majors have been exploring new petroleum reservoirs and evaluating seismic data from field experiments such as well-logging. Roughly speaking about well-logging, boreholes are drilled to the depth to be examined, and logging instruments measure properties of the subsurface inside boreholes with increasing a depth. Therefore, well log data are used as an important source to analyze physical rocks and to obtain corresponding petrophysical properties such as porosity, water saturation, etc. This study uses log data from two wells in the North Sea. Main acquired properties are compressional and shear wave velocities, density, gamma ray, poisson's ratio, and caliper log, as well as estimates of pore fluid content (volume fractions of gas, oil, and water). V_p/V_s ratio and acoustic impedance can be also computed from these basic measured data.

In Fig. 3.1, the depth sample interval is 0.1524m and the total number of measuring points are 14,063 per borehole, and the depth range is from 985.4184m to 3128.4673m. From the total well log data, various kinds of distributions of wave velocities and density show with depth. Among them, the depth range (1,600m$\sim$1,740m) was selected as a representative of upscaling data in this thesis, because this range, unlike other depth ranges, indicates gradually increasing velocity. To confirm how well the upscaling by an MCMC method correctly works, other depth range from 2,610m to 2,730m was studied for upscaling as well.

Figure 3.2 shows increasing compressional and shear velocities and density with depths. In addition, the gamma ray indicates comparatively low values which contains non-shale components. Generally, gamma ray logs measure radioactivity to estimate what kinds of rocks are present in subsurface structures because shales emit lots of gamma rays due to containing more radioactive material. In other words, non-shales such as clean sandstones emit very few gamma rays. These different amounts of gamma radiations distinguish between two rocks. Regarding Poisson ratio, it shows

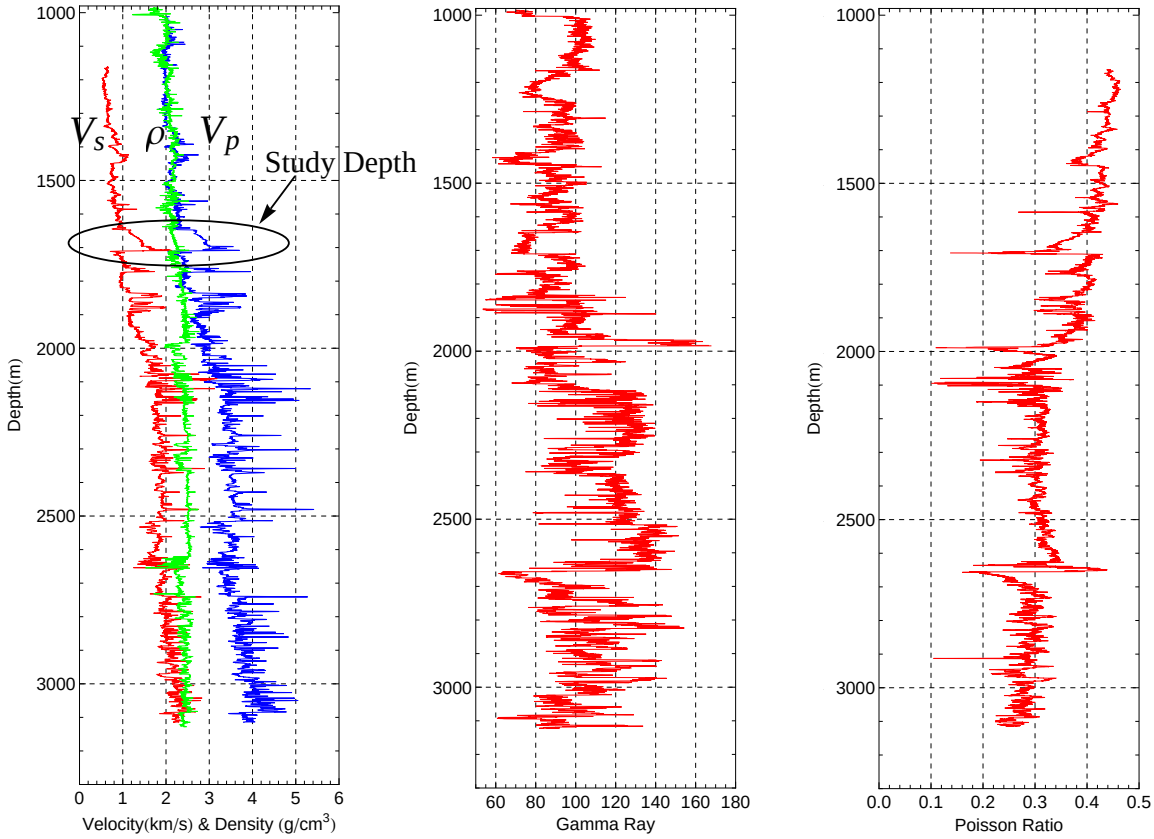


Fig. 3.1. The figures indicate well log data in the total depth range of 985.4184m to 3128.4673m. From the left figure, compressional and shear velocity, density, gamma ray, and Poisson ratio, respectively. As a study depth, the depth range of 1600m to 1740m was selected for upscaling log data.

approximately inverse proportional to wave velocities or density distribution. Usually, Poisson ratio for sandstone is 0.21 to 0.38, and for shale is 0.2 to 0.4, but these values also depend on porosity or water saturation. More porous and water saturated rocks tend to have a little larger values of Poisson ratio.

In practice, it is hard to distinguish sandstone with shale from the Poisson ratio, but we could estimate it closer to sandstone from comparatively higher velocities and lower value of gamma ray with ones of other depth ranges. Therefore, we could presume that this portion might be more porous and water saturated sandstone.

In Fig. 3.3, both Poisson ratio and velocity ratio (V_p/V_s) seem to be equally inverse proportional to velocities (V_p, V_s), but the inverse proportional slopes are considerably different from each other such as -5.585 and -1.109 for compressional velocity.

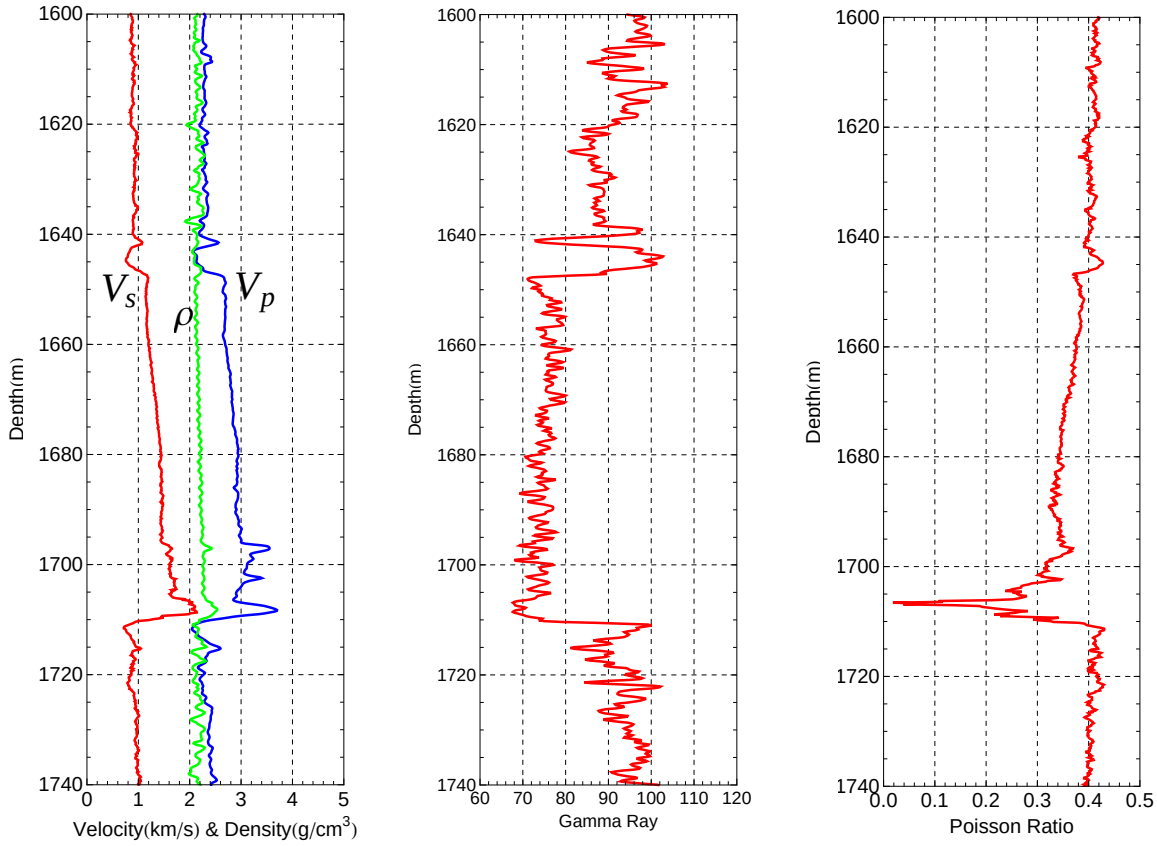


Fig. 3.2. The figures indicate well log data in the chosen depth range of 1600m to 1740m. From the left figure, compressional and shear velocity, density, gamma ray, and Poisson ratio, respectively.

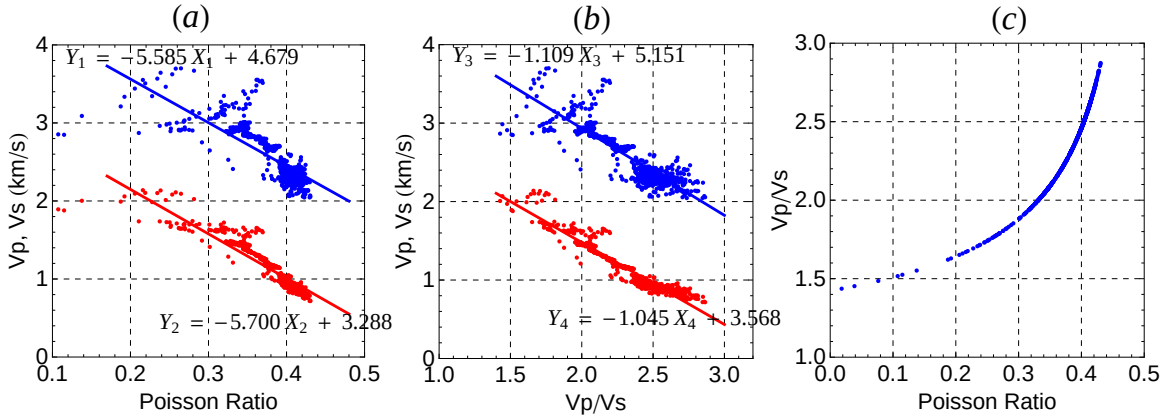


Fig. 3.3. (a) Relationship between Poisson ratio and velocities (V_p, V_s), (b) Relationship between velocity ratio (V_p/V_s) and velocities (V_p, V_s), (c) Relationship between Poisson ratio and velocity ratio (V_p/V_s). The depth range of the well log data is 1600m to 1740m.

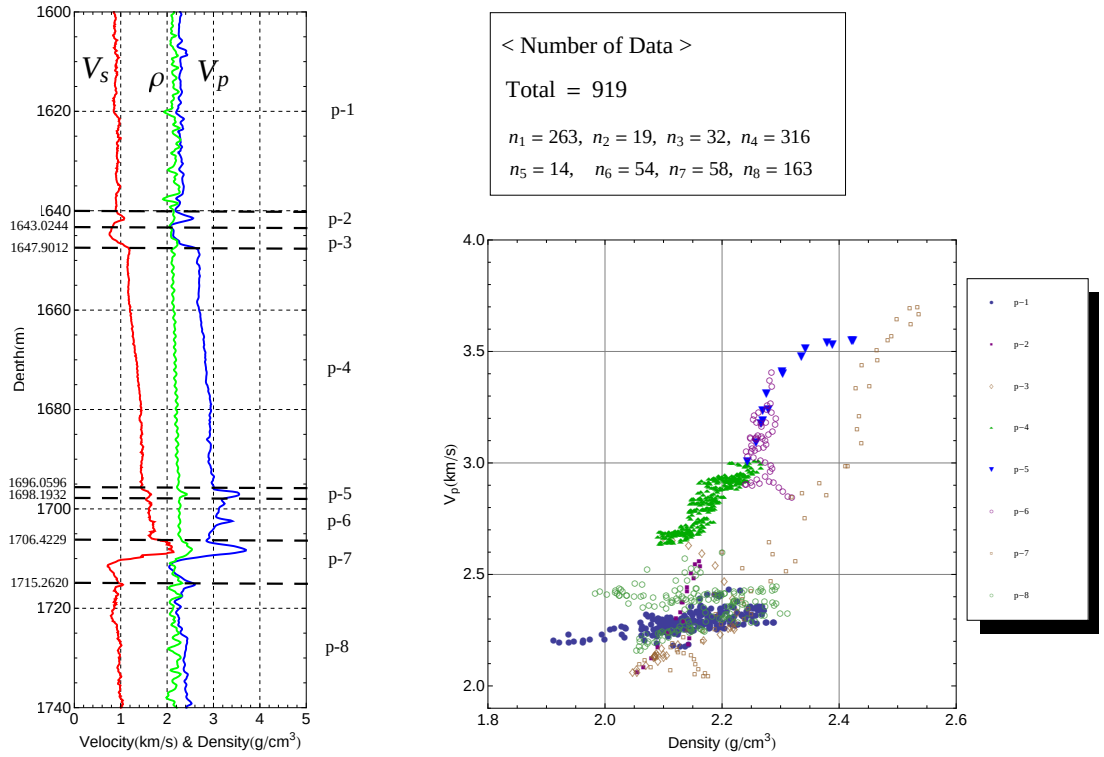


Fig. 3.4. The given depth range is divided by 8 portions showing characteristic distributions. p-1 and p-8 are approximately constant, p-2 and p-5 indicate small unit peaks, and p-4 indicates gradually increasing distribution.

For more detail investigation, the depth range is divided by 8 portions as shown in Fig. 3.4. In each portion, the depth level and the number of data were obtained to represent specific tendencies. The portion 2 and 5 indicate small unit peak, and the portion 4 shows the proportional relationships between densities and compressional velocities. As a rule, densities and velocities are proportionally distributed, but some portions (p-2 and p-6) even show very different distributions. This partition helps us to understand various and uncertain geological characteristics with depth ranges.

Regarding upscaling, its process can be divided largely into two parts. One is the step of defining a model, and the other is the upscaling step of using an MCMC algorithm. In the former, the number of blocked layers and a way of choosing the next boundary are contained in the model, and in the latter, a whole upscaling process is mentioned with an MCMC algorithm based on a Bayesian theory.

3.2 Defining the Model

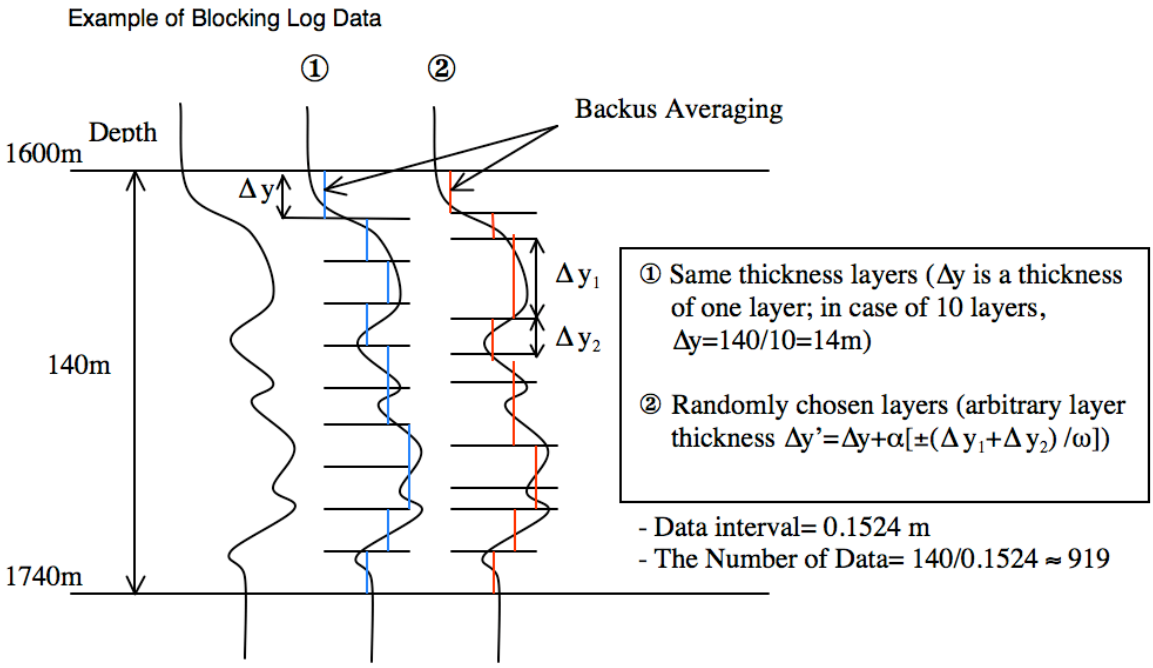


Fig. 3.5. The figure indicates two ways of blocking log data and provides conceptual diagrams about the chosen model (1600m∼1740m). The first way is blocking log data with same thicknesses, and the second one is blocking log data with randomly chosen layers.

With the log data investigated before, defining a model for upscaling is introduced in Fig. 3.5. The MCMC procedure begins by dividing the depth interval into by n layers of equal thickness. An average velocity and a density of each layer are easy to be computed by Backus averaging which well fits to wave properties with large enough wavelengths comparing with the layer thicknesses. Backus proposed an equivalent medium theory in which a heterogeneous medium can be replaced by a homogeneous one containing effectively equivalent elastic constants. In other words, when the ratio of wavelength to the layer thickness is large enough, the wave velocity can be given by an average of the properties of individual layers (Backus, 1962).

$$C_{11}^{*} =< C_{11} > + < \frac{C_{13}}{C_{33}} >^{2} < C_{33}^{-1} >^{-1} - < \frac{C_{13}^{2}}{C_{33}} >,$$

$$C_{12}^{*} =< C_{12} > + < \frac{C_{13}}{C_{33}} >< \frac{C_{23}}{C_{33}} >< C_{33}^{-1} >^{-1} - < \frac{C_{13}C_{23}}{C_{33}} >,$$

$$C_{13}^{*} =< \frac{C_{13}}{C_{33}} >< C_{33}^{-1} >^{-1},$$

$$C_{22}^{*} =< C_{22} > + < \frac{C_{23}}{C_{33}} >^{2} < C_{33}^{-1} >^{-1} - < \frac{C_{23}^{2}}{C_{33}} >,$$

$$C_{23}^{*} =< \frac{C_{13}}{C_{33}} >< C_{33}^{-1} >^{-1},$$

$$C_{33}^{*} =< C_{33}^{-1} >^{-1},$$

$$C_{44}^{*} =< C_{44}^{-1} >^{-1},$$

$$C_{55}^{*} =< C_{55}^{-1} >^{-1},$$

$$C_{66}^{*} =< C_{66} >,$$

where C_{ij} (i,j = 1,2,3,...,6) represent the effective elastic constants in two-index Voigt notation and C_{ij}^{*} represent the effective elastic constants over the scale of the averaging window length. The bracket <> is an integral over the size of the window (Liner and Fei, 2007).

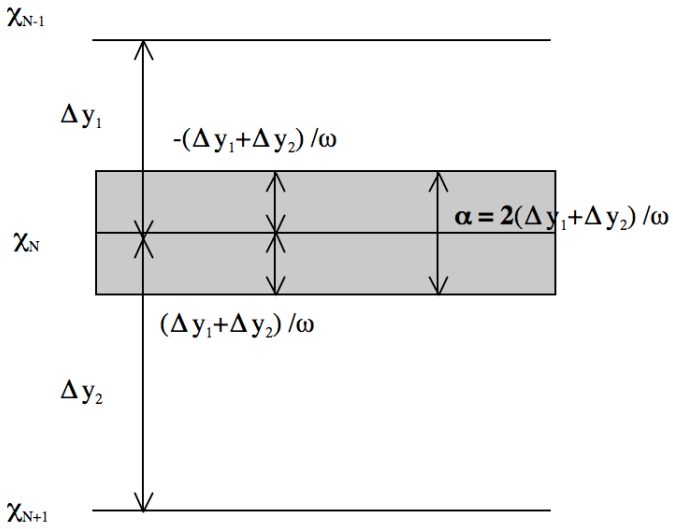


Fig. 3.6. The figure shows a way of choosing the next boundary level. In an arbitrary layer boundary, the line χ_N moves within the range of α, and the width of random walk depends on the thickness of upper and lower layers. In this figure, ω indicates a coefficient of step size as used in the MCMC algorithm.

As the next step of choosing a new model in each iteration, a way of getting a new layer boundary is described in Fig. 3.6. In an arbitrary layer boundary, an arbitrary line χ_N is chosen by an MCMC algorithm within the gray area which depends on the thickness of the upper and lower layers. A way of selecting the next model is now finished, and then move to a main procedure for upscaling log data.

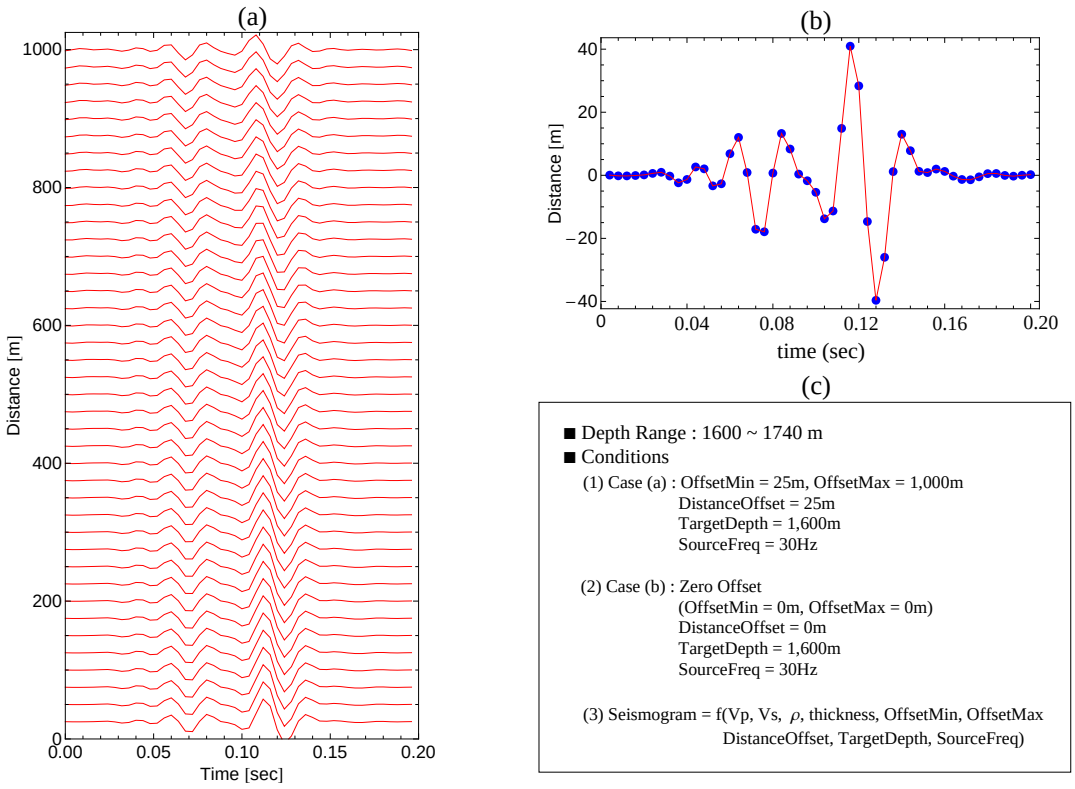


Fig. 3.7. (a) The figure shows a general seismogram calculated from the measured data (1600m∼1740m). (b) The figure indicates a seismogram in case of zero-offset. (c) Basic conditions to get the figures of (a) and (b) are represented.

In Fig. 3.7, synthetic seismograms are calculated from the measured data with the given conditions (c), and seismograms are a function of wave velocities (V_p, V_s), density, thickness of layers, minimum offset, maximum offset, and target depth. As typical input values, the minimum offset of 25m, the maximum offset of 1000m, the offset interval of 25m, and the source frequency of 30Hz were used in this study. The figure (b) shows a zero-offset seismogram with which we can calculate RMS error values by comparing with a reference seismogram.

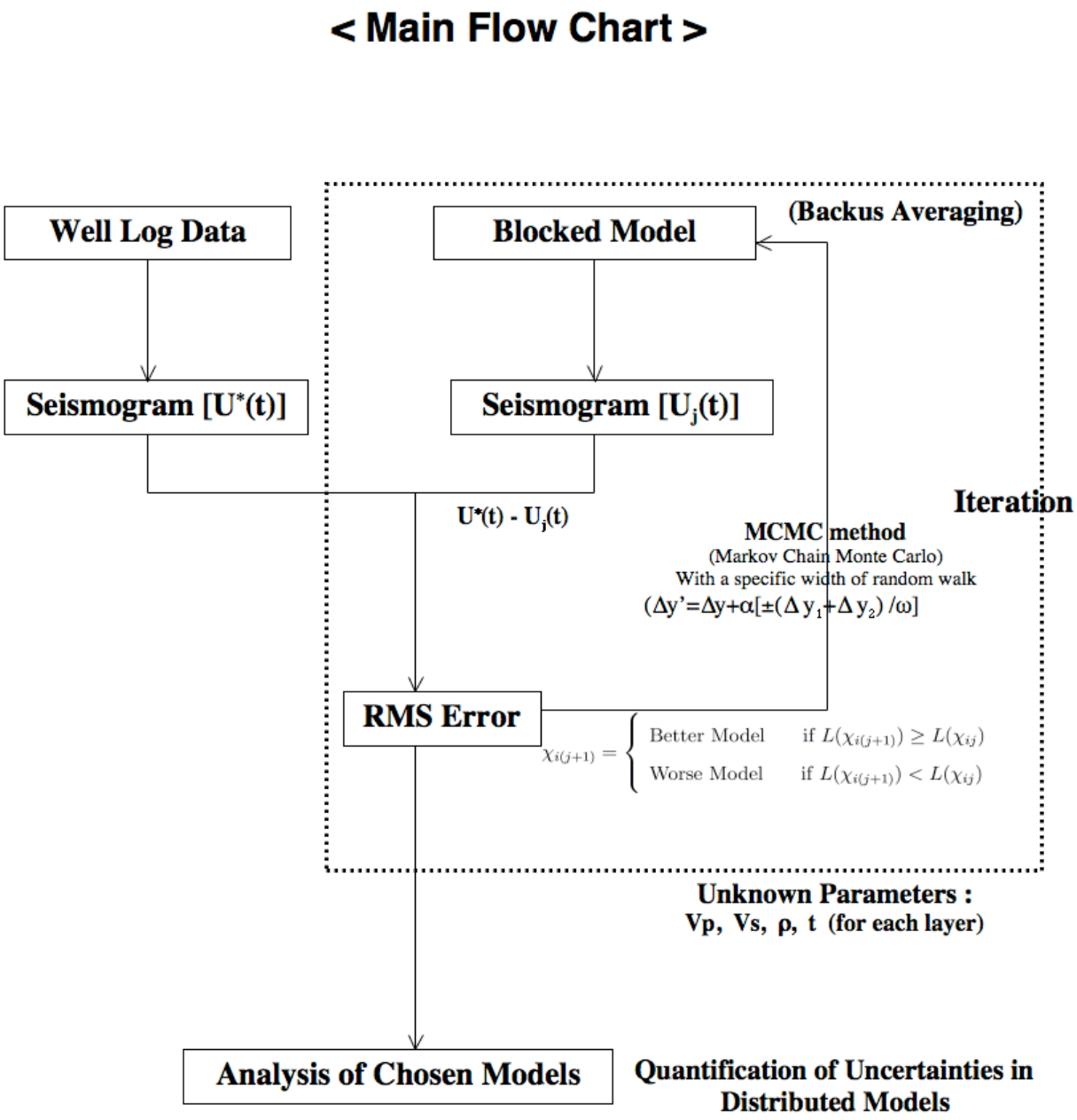


Fig. 3.8. This flow chart indicates a schematic procedure for upscaling by an MCMC method. Ultimately, the chosen models are used to quantify uncertainties in upscaling log data.

Fig.3.8 shows an overall flow chart of the main algorithm. Which model is better here can be verified by comparison of the reference seismogram with the calculated synthetic seismogram. To be specific, RMS error values could be good indicators of how well established the algorithm is, because they represent a degree of similarity with the reference seismogram. If a seismogram from upscaled data does not fit to the one from well log data, RMS error value gets larger, and the algorithm has less confidence. The other verification of the algorithm comes from the overlapped figure of two unit seismograms because well matched seismograms could be considered as originating from very similar data sets. Based on this main procedure for the algorithm, the uncertainties in parameters could be quantified with statistical analyses of reasonably chosen models.

To begin a specific procedure for upscaling, we need to compute a reflection seismogram ($U^*(t)$) for all measured data as shown in Fig. 3.8, which will be continuously used as reference data during whole iteration, and then a blocked model is chosen according to the way explained before. With this model, we can compute a new seismogram ($U_j(t)$). From the measured seismogram ($U^*(t)$) and the calculated seismogram ($U_j(t)$), RMS error values are computed and applied to the likelihood ($L(\chi_{ij})$) and the acceptance probability (P_{accept}). Finally, we should make a decision to keep or reject new models, and repeat from the step of generating new models to the step of determining to keep or reject new models during the whole iteration. Depending on a σ value, the number of layers, iteration, and a coefficient of step size, distributions of new models could be very different.

3.3 Choice of a Reasonable σ Value in Actual Upscaling

Before choosing a suitable σ value, we need to consider two important points in an upscaling procedure. One is that a range of square RMS error values should be small enough to get reliable models, because a large range of error values literally means sometimes taking models that are very different from the reference seismogram. The other is that the number of newly taken models should be large enough to analyze statistically because it is difficult to quantify uncertainties in upscaling log data with a restricted number of models. In this experiment, σ values of 1, 3, and 5 are chosen first because the range of $\triangle S^2$ values shows approximately from zero to 200, and the ratio ($\frac{\triangle S^2}{2\sigma^2}$) shows reasonable ranges with these σ values as shown in Fig. 2.1.

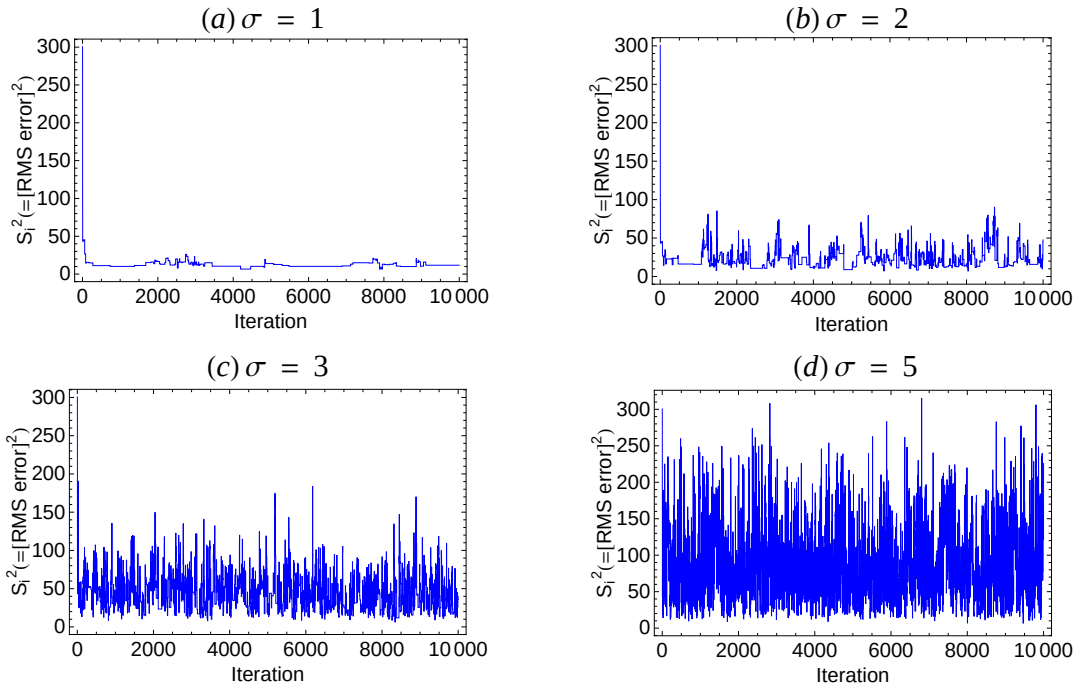


Fig. 3.9. The figures show the distributions of square RMS error values depending on σ values of 1,2,3, and 5. The initial parts of each figure indicate very high values which need to be corrected.

In Fig. 3.9, (a) (σ=1) shows that very few models are selected although most square RMS error values are very small, and (d) (σ=5) indicates the very large range of square RMS error values, and a large number of models are taken during the iteration, whereas (c) (σ=3) seems to have a sufficient number of models and the reasonable range of square RMS error values. However, if we have a closer look at it, (c) also has a little large range of error values comparing with the one of (b) (σ=2). Therefore, (b) among the four cases is yielding most reasonable models with satisfying two conditions mentioned before. In the figure 3.9, the initial parts of each figure need to be removed because the first 100 iteration shows very large error values comparing to other iteration and do not show the important range of values where the Markov Chain samples search models.

Fig. 3.10 describes main sequential results of an MCMC algorithm. First, square RMS errors (S_i^2) can be calculated from taken models. With these RMS error values, the likelihood ($L(\chi_{ij})$) can be described as an exponential function.

$$L(\chi_{ij}) = k\ exp[-\frac{1}{2}\frac{S^2(\chi_{ij})}{\sigma^2}] \tag{3.1}$$

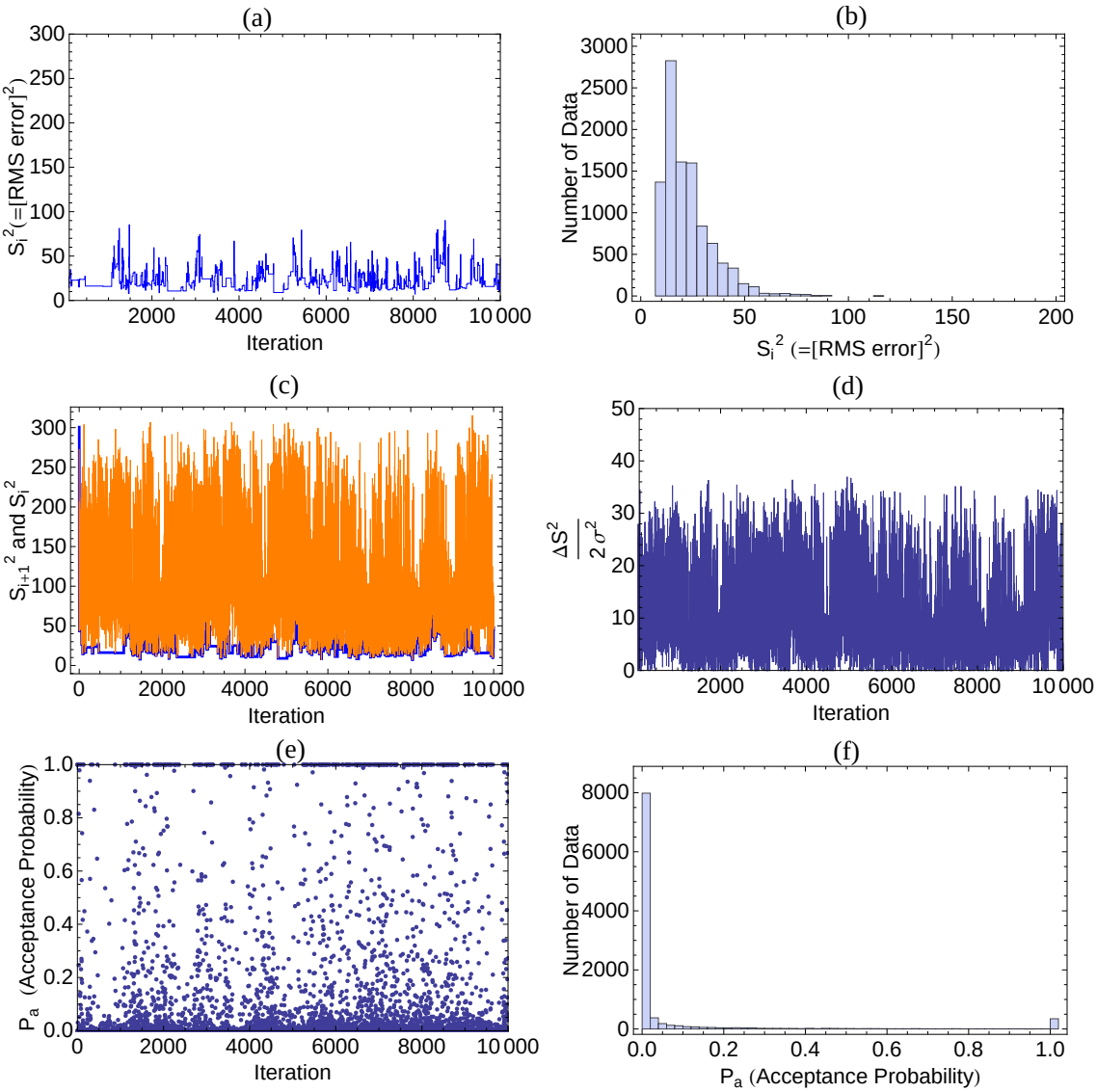


Fig. 3.10. The figures show the main results of the MCMC algorithm with input variables (σ=2, iteration=1,000,000, the number of layers=10, a coefficient of step size=1). (a) Distribution of square RMS error values with iteration, (b) Histogram of square RMS error values, (c) Distribution of square RMS error value in both previous models (blue) and newly generated models (orange) with iteration, (d) Distribution of the ratio of $\frac{\triangle S^2}{2\sigma^2}$ with iteration, (e) Distribution of the acceptance probabilities with iteration, (f) A Histogram of the acceptance probabilities.

The likelihood is used to determine which models are better or worse as below.

$$\chi_{i(j+1)} = \begin{cases} \text{Better Model} & \text{if } L(\chi_{i(j+1)}) \geq L(\chi_{ij}) \\ \text{Worse Model} & \text{if } L(\chi_{i(j+1)}) < L(\chi_{ij}) \end{cases} \tag{3.2}$$

However, all worse models are not rejected, but some worse ones are accepted within an allowable probability as following.

$$P_{accept} = \begin{cases} 1 & \text{if} \quad L(\chi_{i(j+1)}) \geq L(\chi_{ij}) \\ \exp(-\frac{\triangle S^2}{2\sigma^2}) & \text{if} \quad L(\chi_{i(j+1)}) < L(\chi_{ij}), \end{cases} \tag{3.3}$$

where $\triangle S^2 = S^2(\chi_{i(j+1)}) - S^2(\chi_{ij})$.

After the iteration, we can get a number of reasonable models, 99,000 in this case, and the uncertainties in the models are quantified through statistical analyses such as histograms and PDFs of the distributed models. The ratio ($\frac{\triangle S^2}{2\sigma^2}$) is obtained from the difference between two successive square RMS errors ($\triangle S^2$) and the given σ value. Therefore, the acceptance probabilities such as (e) and (f) in Fig. 3.10 can be determined from the relationship between the acceptance probability and the ratio of $\frac{\triangle S^2}{2\sigma^2}$ (see Fig. 2.1). With the determined models, we can obtain model parameters (V_p, V_s, ρ, thickness) as well.

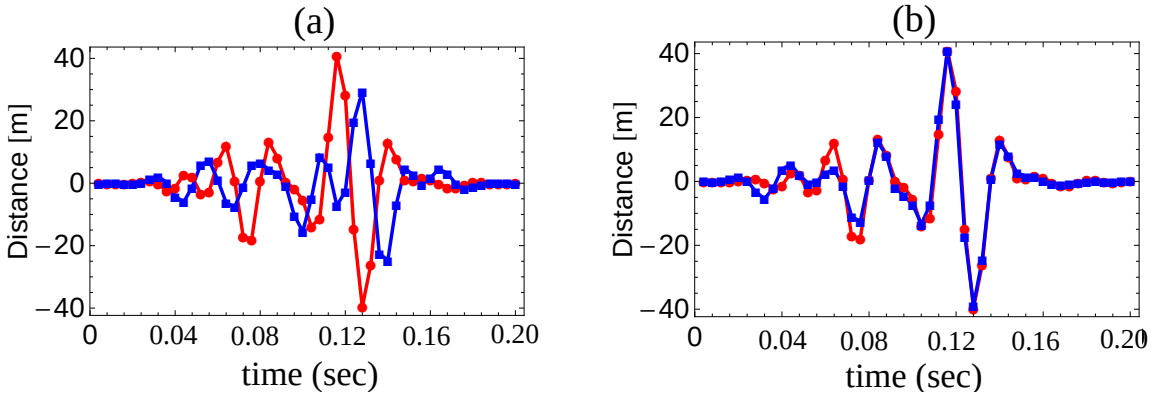


Fig. 3.11. (a) The figure shows the initial comparison of the reference seismogram (red) with the first seismogram (blue) from the initial model. The initial RMS error value is 17.34. (b) The figure indicates the comparison of the reference seismogram (red) with the calculated seismogram (blue) of the best model after the iteration. The RMS error value is 2.61.

3.4 Feasibility of an MCMC Method in Upscaling

An MCMC method can produce very reliable results in upscaling log data because initial large square RMS error values can be significantly reduced as optimal interface depths are determined. Fig. 3.11 represents well matched seismograms which consist of the reference seismogram from well log data and the calculated seismogram from blocked model parameters. The seismograms are calculated with conditions of zero-offset and the given input variables (the number of layers=10, iteration=10,000, a coefficient of step size =1, and σ=2). Two seismograms are overlain to distinguish the difference graphically. Figure (a) shows a big difference between the reference seismogram (red line) and the calculated seismogram (blue line) in the initial model. However, this disparity changed into a considerable similarity such as the figure (b), which means an MCMC method could provide significantly credible results for upscaling. Even though some points are not exactly accord with each other, this difference is small comparing with the large range of synthetic data (-40$\sim$40) in seismograms. In practice, the range of RMS error values after one million iteration is the minimum 2.61 to the maximum 9.49, and the average value for all taken models is 4.73. Furthermore, these small RMS error values could be even more reduced by controlling input variables such as a σ value, iteration, the number of layers, and a coefficient of step size.

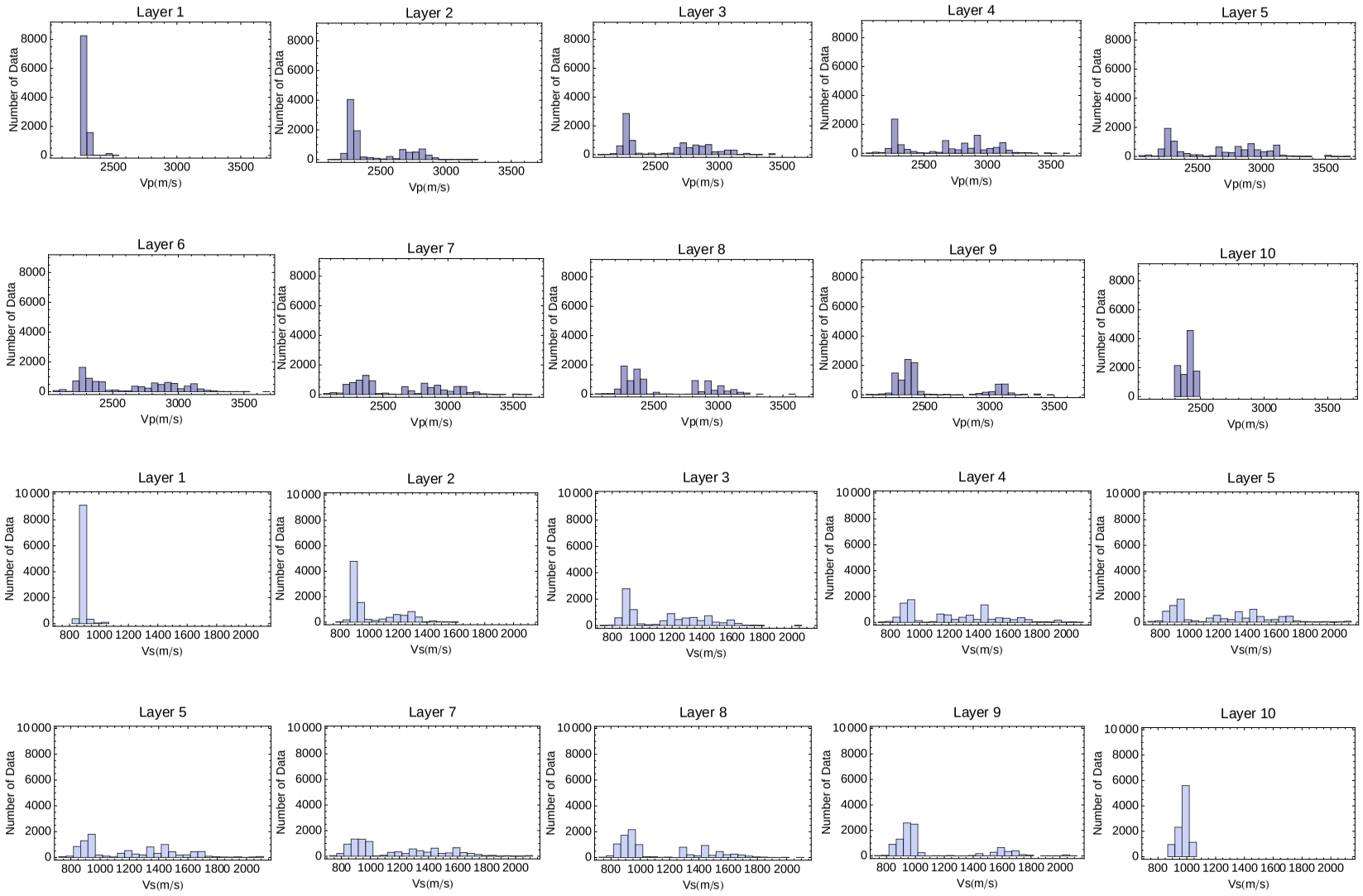


Fig. 3.12. Histograms of V_p and V_s from the layer 1 to the layer 10 during 10,000 iteration.

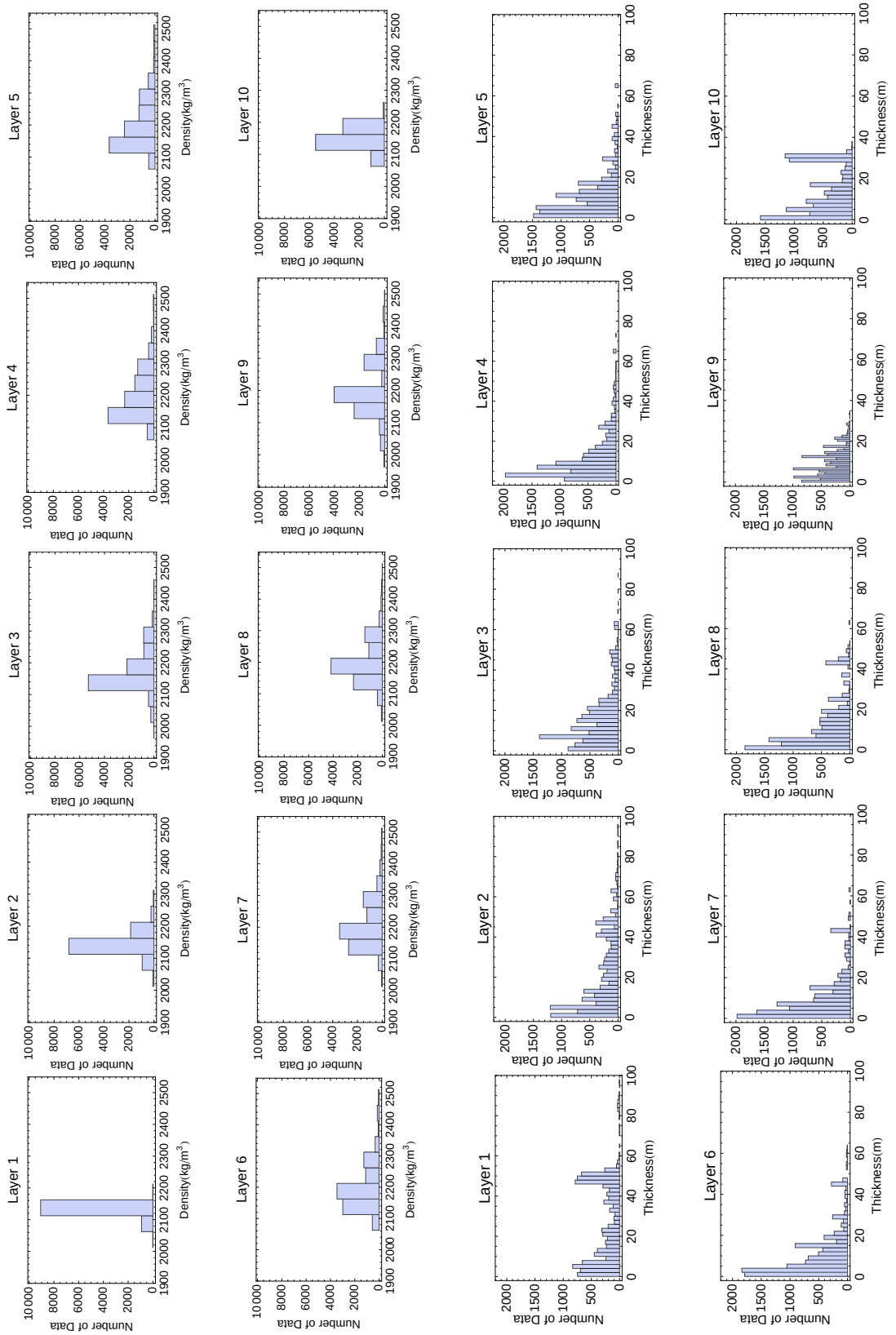


Fig. 3.13. Histograms of density and thickness from layer 1 to layer 10 during 10,000 iteration.

Fig. 3.12 and 3.13 indicate histograms of parameters (V_p, V_s, ρ, thickness) in each layer. As known in Fig. 3.2, histograms of layer 1 and 10 show characteristics concentrated on specific depth ranges, and ones of layer 2 to 9 indicate comparatively wide distribution because wave velocity or density shows gradually increasing values in the mid depth ranges. Therefore, histograms of model parameters could vary depending on chosen depth ranges.

3.5 Control Variables on Upscaling in an MCMC Algorithm

The MCMC algorithm in chapter II has four main input variables such as a σ value, iteration, the number of layers, and a coefficient of step size. These variables seem to be independent of each other but systematically connected to mutual functions. For this reason, we need to examine the influences of input variables on generating reliable models. Better understanding the influence of each variable could make it easy to effectively control the generating system of models, and quantification of uncertainties in upscaling could be correctly concluded on condition of satisfying this premise.

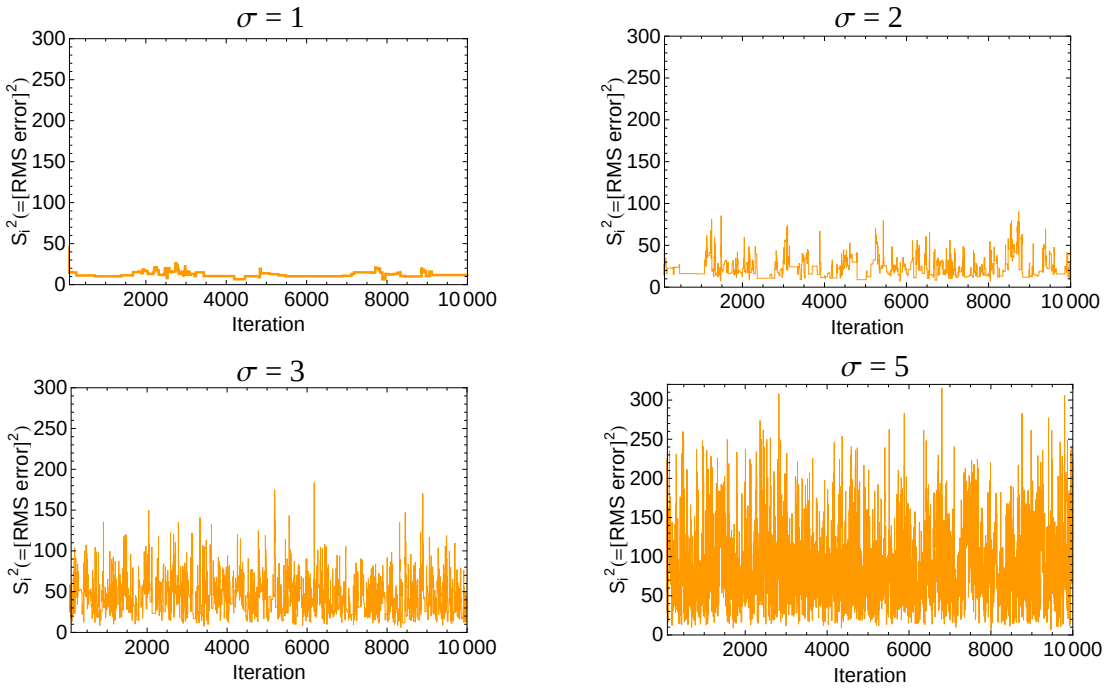


Fig. 3.14. The figures show the distributions of square RMS error values depending on σ values. Used σ values are 1, 2, 3, and 5, respectively.

3.5.1 Influence of a σ value

As mentioned in the section 2.2, the σ value (see the equation 2.3 and 2.4) is the most influential factor to determine probabilities taking new models. Fig. 3.14 shows distributions of square RMS error values (S^2) depending on a σ value. A larger σ value causes a wider variation and a more rapidly fluctuating distribution of square RMS error values. The initial models (1∼100) are not displayed to simplify statistical analyses.

Table 3.1. Results of square RMS error values depending on σ values.

	σ=1	σ=2	σ=3	σ=5
Max(S^2)	23.0	90.1	183.9	315.6
Min(S^2)	6.6	6.8	6.2	6.8
Avg(S^2)	12.1	22.4	41.3	77.5
Number of models	96	706	2,176	5,284

Table 3.2. Comparative analysis of other square RMS error values normalizing by values at $\sigma = 1$ (see Table 3.1)

	σ=1	σ=2	σ=3	σ=5
Max(S^2)	1	3.92	8.00	13.72
Min(S^2)	1	1.03	0.94	1.03
Avg(S^2)	1	1.85	3.41	6.40
Number of models	1	7.35	22.67	55.04

More specifically speaking, the average values of S^2 increased from 12.1 to 22.4, 41,3, and 77.5, respectively (see Table 3.1). In addition, the number of models that were retained according to the selection rule in equation 3.3 significantly increased from 96 to 706, 2,176, and 5,284, respectively (see Fig. 3.15). For a better comparison, Table 3.2 provides an evident tendency of how many times S^2 values increased comparing with S^2 values at $\sigma = 1$. We need to notice that maximum and average

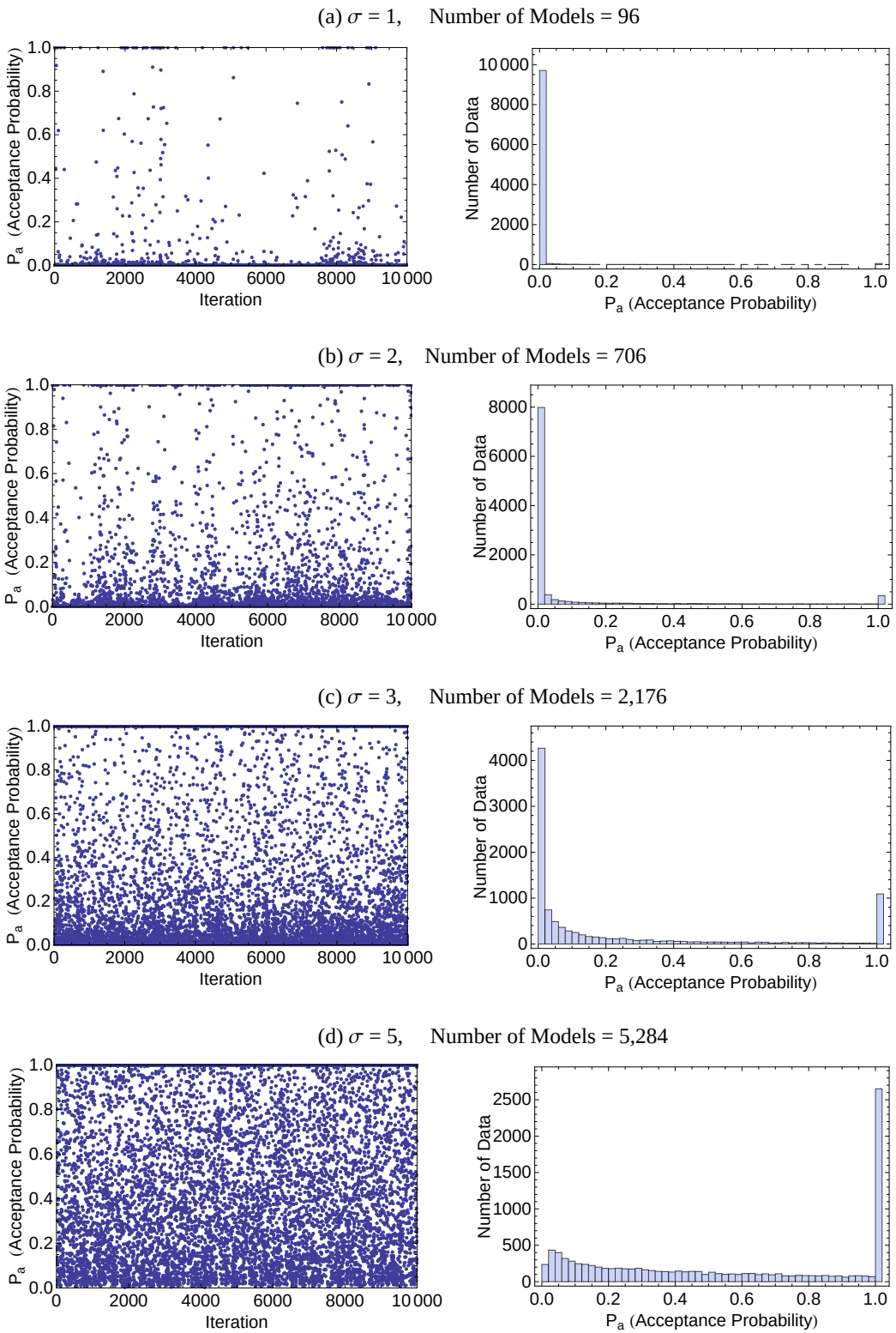


Fig. 3.15. The figures show different acceptance probabilities depending on σ values (a) The number of models is 96 in the σ value of 1. (b) The number of models is 706 in the σ value of 2. (c) The number of models is 2,176 in the σ value of 3. (d) The number of models is 5,284 in the σ value of 5.

values of S^2 increased considerably, but minimum values barely changed depending on σ values. Particularly, the number of models shows the fastest increase, which means that even a small change of a σ value had quite a little influence on the number of models.

3.5.2 Influence of Iteration

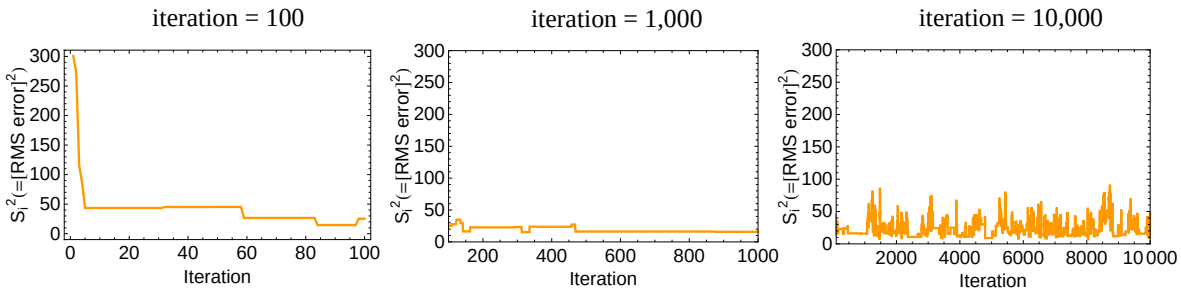


Fig. 3.16. The figures show the distributions of square RMS error values depending on iteration. Iteration is 100, 1,000, and 10,000, respectively. Initial models (1∼100) are not displayed to simplify statistical analyses.

Table 3.3. Results of square RMS error values depending on iteration.

	iteration=100	iteration=1,000	iteration=10,000
Max(S^2)	300.7	34.7	90.1
Min(S^2)	14.6	15.3	6.8
Avg(S^2)	41.2	19.0	22.4
Number of models	10	22	706

As shown in Fig. 3.16, more iteration could result in selecting larger number of models and smaller values of square RMS error. If, however, a σ value is small enough, such as below 1, the frequency of random walk in MCMC is very few, and almost all models are rejected. In this case, increasing iteration is not useful any more. Therefore, input variables should be considered at the same time when simulating an MCMC algorithm. Table 3.3 displays summary of error values obtained in each case, as total numbers of iterations change.

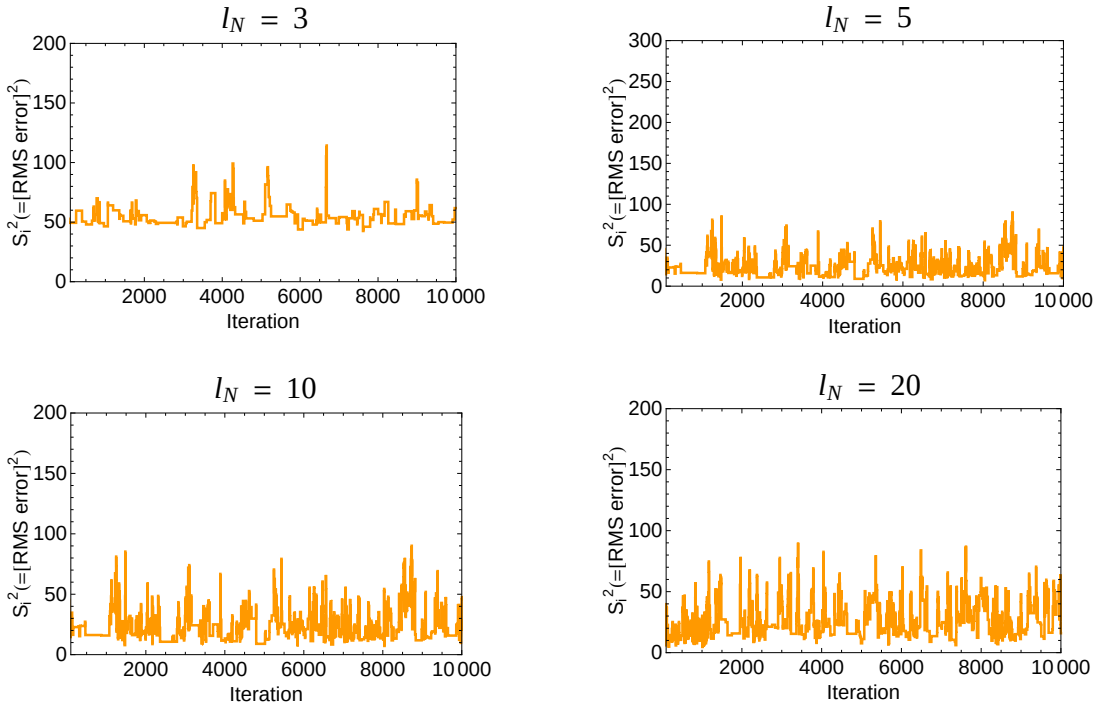


Fig. 3.17. The figures show the distributions of square RMS error values depending on the number of layers. The number of layers is 3, 5, 10, and 20, respectively. Initial models (1∼100) are not displayed to simplify statistical analyses.

Table 3.4. Results of square RMS error values depending on the number of layers.

	l_N=3	l_N=5	l_N=10	l_N=20
Max(S^2)	114.3	90.0	90.1	90.2
Min(S^2)	42.16	12.0	6.8	3.7
Avg(S^2)	54.6	28.0	22.4	24.5
Number of models	168	230	706	864

3.5.3 Influence of the Number of Layers

The Table 3.4 and Fig. 3.17 represent two characteristic things. One is that maximum values of S^2 do hardly change in the number of layers of 5, 10, and 20. The other thing is that average values of S^2 in a case of 10 layers is even a little smaller than one of the case of 20 layers. It is natural that more blocked layers have the smaller S^2 values and more models, but our interest in upscaling is in choosing fewer layers with smaller error values. Therefore, we need to choose the number of layers as small as possible with considering both square RMS error values and the number of models.

3.5.4 Influence of a Coefficient of Step Size

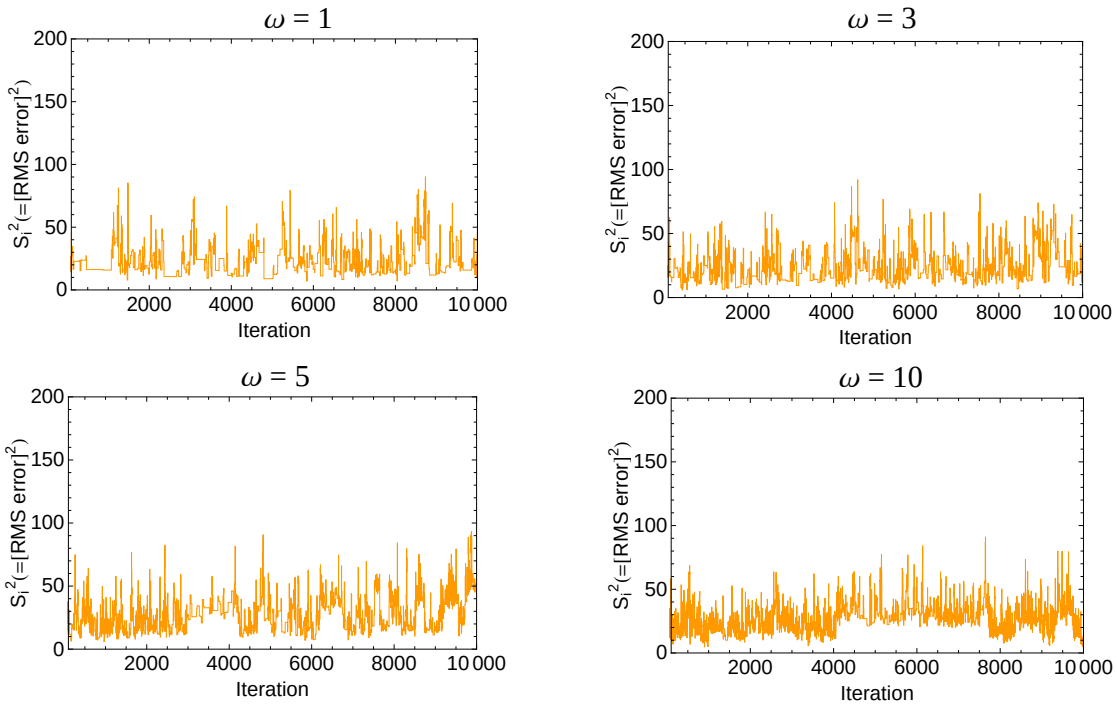


Fig. 3.18. The figures show the distributions of square RMS error values depending on a coefficient of step size in the MCMC algorithm. A used coefficient of step size is here 3, 5, 10, and 20. Initial models (1∼100) are not displayed to simplify statistical analyses.

Fig. 3.18 displays the development of error values with iteration for different values of ω. In Table 3.5, maximum, minimum, and average of S^2 show very few change comparing with other input variables. Therefore, a coefficient of step size

Table 3.5. Results of square RMS error values depending on a coefficient of step size.

	ω=1	ω=3	ω=5	ω=10
Max(S^2)	90.1	92.0	93.2	90.7
Min(S^2)	6.8	5.7	5.7	4.6
Avg(S^2)	22.4	22.9	26.7	26.5
Number of models	706	1,234	1,645	2,675

is comparatively less important factor to get new models. However, we also need to notice that the number of models increases depending on the larger value of a coefficient of step size (ω). Because a larger ω value means a smaller step size from the range ($-\frac{\chi_{(i-1)j}+\chi_{(i+1)j}}{\omega} \sim \frac{\chi_{(i-1)j}+\chi_{(i+1)j}}{\omega}$) as mentioned in the MCMC algorithm, the result of larger ω value naturally shows the larger number of models. With these characteristics, we could apply a ω value to the case of requiring many good models.

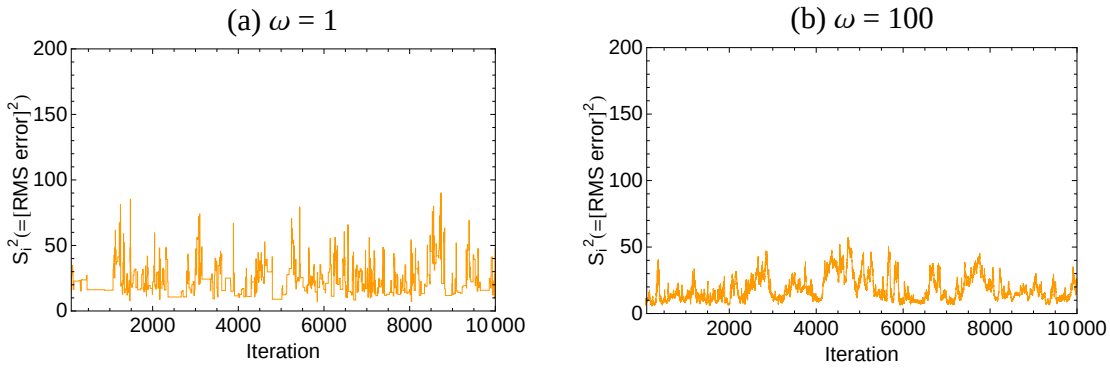


Fig. 3.19. The figures show the results of another way of being able too take many reasonable models. The model of the minimum square RMS error value in the first figure ($\omega = 1$) was used as an initial model in the second figure ($\omega = 100$). Initial models (1∼100) are not displayed to simplify statistical analyses.

In Fig. 3.19, (a) shows a general result of square RMS error. With the results, we can get the model with the minimum S^2 of 6.8. It is assumed to be a good model because the model has a very small error value, but it could be accidentally happened by generating random numbers. Therefore, an important element having more confidence in models is that lots of models should be produced within a small range of error values. To get more reliable models, it is recommended that the minimum error

Table 3.6. Comparison of two results of (a) and (b) in Fig. 3.19.

	ω=1	ω=100
Max(S^2)	90.1	57.4
Min(S^2)	6.8	5.7
Avg(S^2)	22.4	18.7
Number of models	706	9,216

value in Fig. 3.19 (a) be applied to the initial model of the figure (b). From (b), many credible models could be produced because the step size of (b) is considerably small. The average value of S^2 in (b) is 18.7 which means a decrease of 16.5 percent comparing with the one of (a). Above all, the number of models in (b) is 9,216 which means an increase of approximate 12 times comparing with the one of (a). This suggested way could simultaneously satisfy two conditions consisting of taking the large enough number of models and obtaining reasonable models with small enough error values. Acquisition of many models with small error values facilitates quantification of uncertainties in upscaling log data in a statistical way.

3.6 Statistical Quantification of Uncertainties

Fig. 3.20 is a total histogram for interface depths of all layers for all models selected by the MCMC algorithm. The histogram shows that a large percentage of models have interfaces near depths of 1050m, 1695m, and 1710m. In particular, the number of chosen interface depths around 1,710m is almost same as the number of iterations, which means most models include that the depth as an interface. In the acoustic impedance of Fig. 3.21, this depth indicates the largest variation among the total acoustic impedance, whereas the depths around 1,622m display very small number of samples compared with other depths because the small variation of acoustic impedance shows around this depth. In other words, relatively small contrast in the variation of the acoustic impedance means that the reflection coefficient in the depth is smaller than the ones of other depths showing larger histograms. In the well log data (see Fig. 3.2), it was very hard to pick layer levels for upscaling, but we

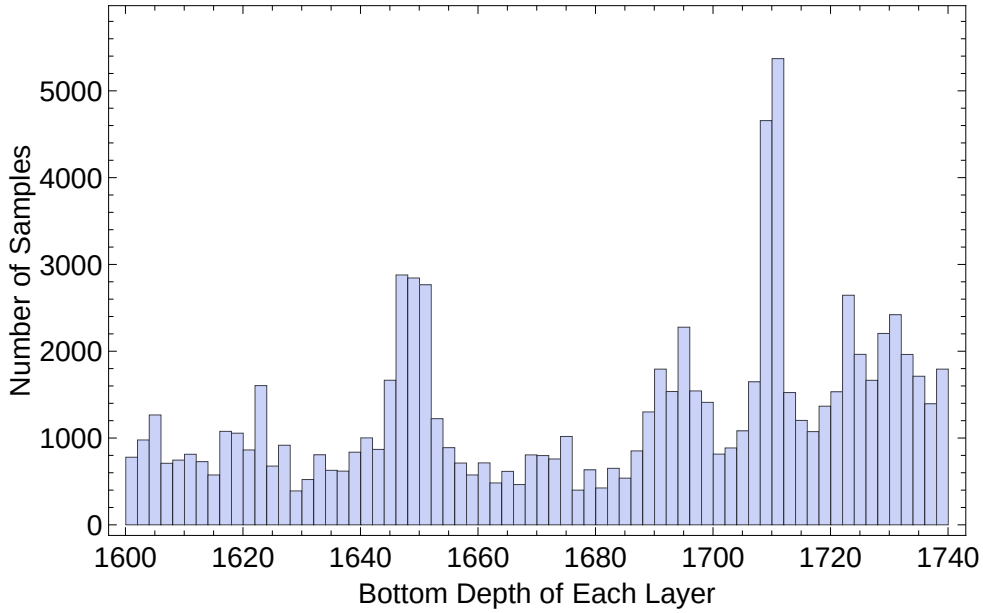


Fig. 3.20. The figure shows the total histogram of interface depths of all chosen layers during the iteration. Approximately, three portions nearby 1,650m, 1,695m, and 1,710m were more often taken as the interface levels of the blocked layers. The number of the interface levels is nine in each model. In case of 10,000 iteration, therefore, the total number of interface levels becomes 90,000 as shown in this histogram.

could give a great confidence in choosing those depths after upscaling by an MCMC method. Furthermore, the histogram provides good information to determine how many and which levels need to select. That is, it is possible to quantitatively analyze uncertainties in upscaling log data.

With the most frequently chosen models, we can estimate reasonable interface depths of blocked layers. Fig. 3.21 compares original fine-scale logs with a blocked model chosen by selecting the model that was unchanged for the largest number of iterations. In other words, the Metropolis selection rule (equation 3.3) rejected a large number of changes, indicating that the misfit compared to the reference seismogram was small and that it is a good fit. Comparing the depths of the interfaces to the distribution in Fig. 3.20 shows that the model has boundaries consistent with the peaks in the distribution.

Depending on input variables, upscaling is automatically fulfilled independent of depth ranges or perturbation of well log data, which makes an MCMC algorithm more reliable because there is no artificial manipulation. The blocked lines in Fig. 3.21 also could be considered as the one of representatives close to the actual log data

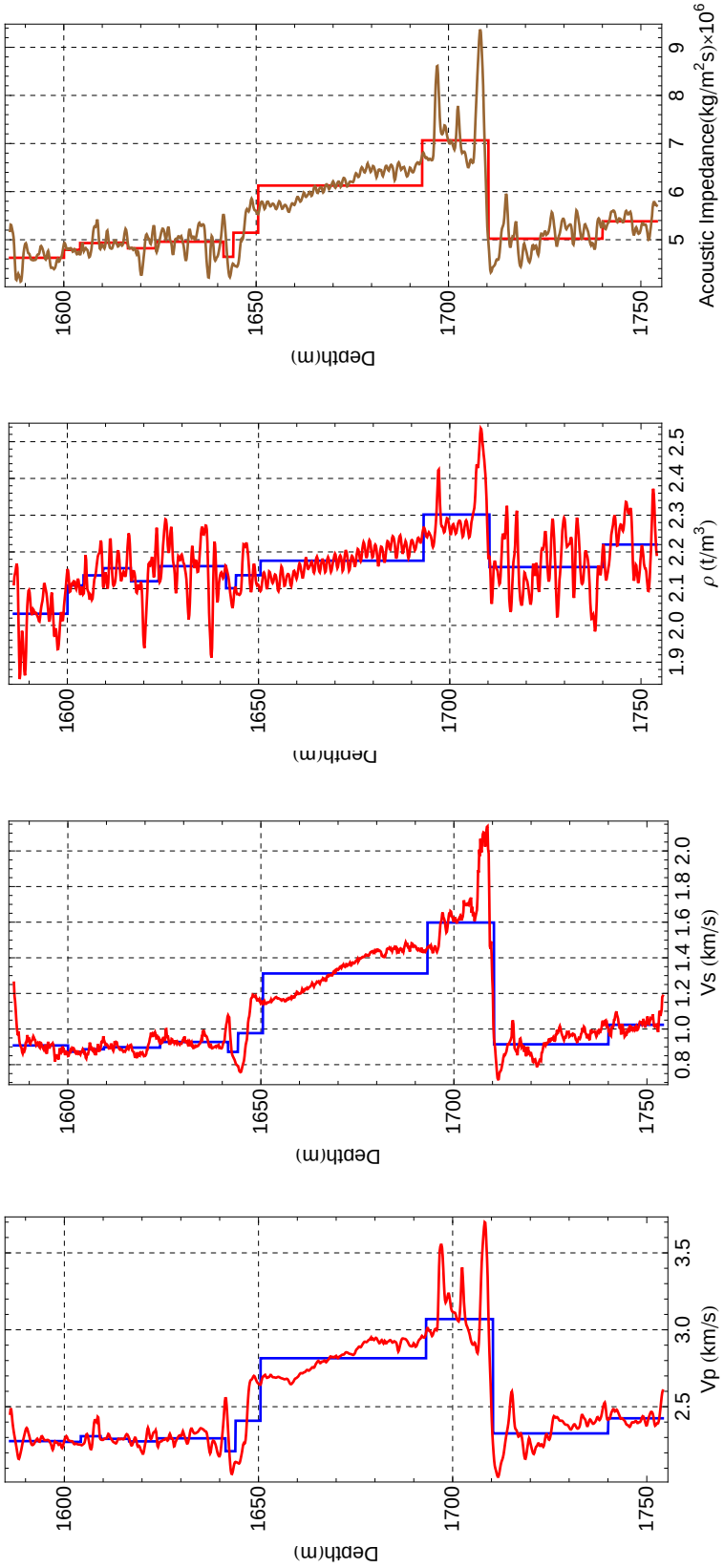


Fig. 3.21. The figures indicate the results of upscaling by an MCMC method. The solid line shows the sample model described in the text (iteration=10,000, the number of layers=10, σ=2, and a coefficient of step size=1).

because many other models with the same RMS error values might exist. Another interesting experiment was made with five layers as an input variable.

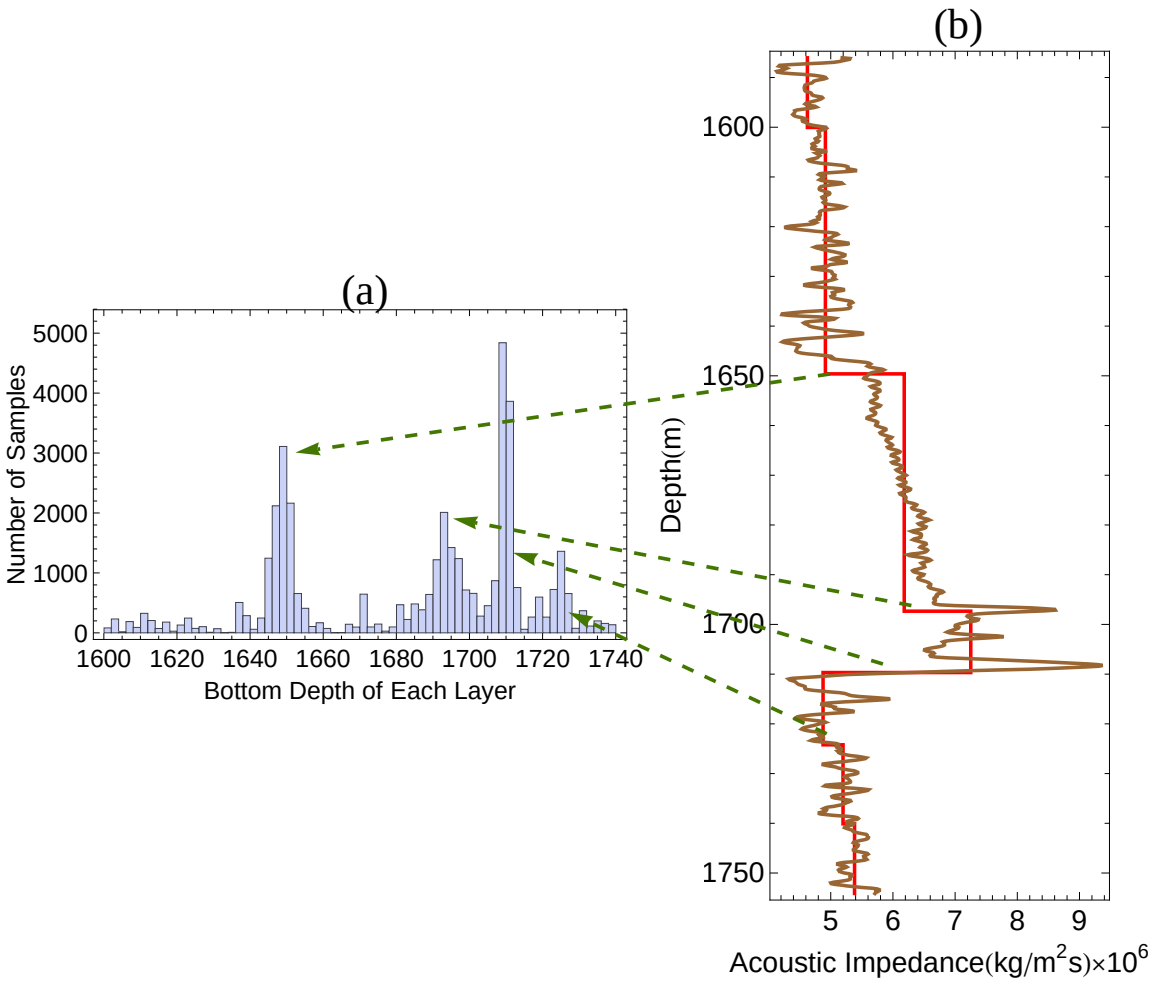


Fig. 3.22. When five layers were used as an input variable ($l_n = 5$), the figure (a) shows the total histogram of interface levels, and the solid line of the figure (b) indicates the most frequently chosen model. Comparing with the Fig. 3.20, more obvious histogram appears in nearby the four portions (1,650m, 1,695m, 1,710m, and 1,725m).

Fig. 3.22 shows a histogram and a reliable blocked model. Comparing with the Fig. 3.20, the histogram appears more obvious and gives us a more confident result. Even a little ambiguous peaks below a depth of 1,720m indicated more clear-cut depths. From the histogram, we can estimate that approximately five blocked layers can be selected for upscaling. This information is very useful as it is, because we do not have any good information at first from the well log data. With these results, we

could understand that five layers are enough to give a good estimation of upscaling well log data.

3.7 Application to Other Depth Layer with Hydrocarbon

Up to now, we have investigated how to make upscaling well log data based on an MCMC method, and what main factors make effects on acquisition of reasonable models. In addition, uncertainties in upscaling log data could be quantitatively analyzed by using statistical approach. With these useful results, we need to confirm how well an upscaling method can be applied to other depth ranges.

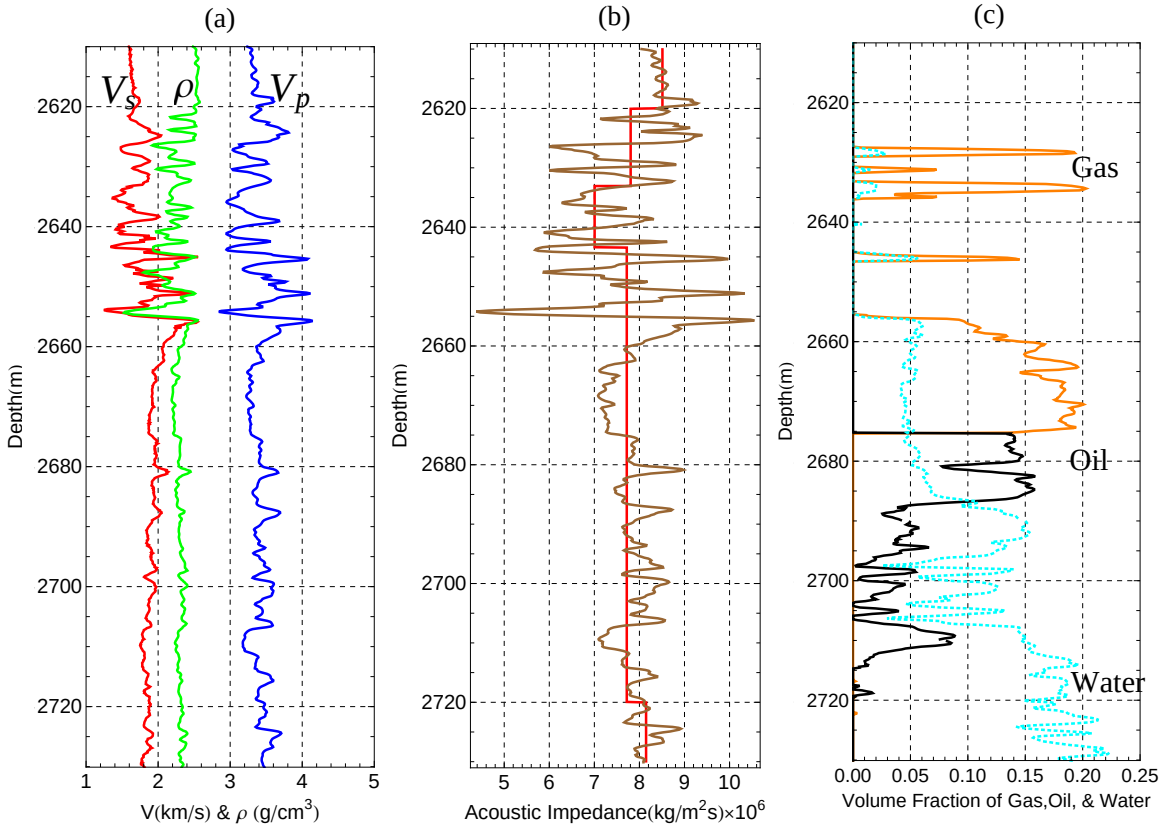


Fig. 3.23. (a) The figure shows the distributions of V_p, V_s, and density in another depth range (2,610m∼2,730m). (b) The figure represents upscaling the acoustic impedance in the same depth range. (c) The figure indicates the volume fractions of gas (orange), oil (black), and water (cyan).

Fig. 3.23 shows a set of log data from a different depth interval, which is of interest because it is a hydrocarbon reservoir. In this experiment, input variables

were used with the conditions that the number of layers is 3, the σ value is 1, the iteration is 10,000, and the coefficient of step size is 1, respectively. The figure also indicates a rapidly varying distribution of acoustic impedance in the initial portions. Therefore, these portions are more probably chosen as a bottom level for upscaling than other portions. In (b) of Fig. 3.23, the red solid line indicates the one good model obtained from the mode of the many selected models. Although the depths around 2650m showing a large contrast of acoustic impedance was not picked for upscaling, more irregular portions, instead, were selected to get well matched seismograms with the reference synthetic seismogram. In case of the symmetry of a right and left side in the distribution of well logs, this algorithm tends to less pick those depths because the symmetric distribution of well logs does not have a large effect on constructing a seismogram.

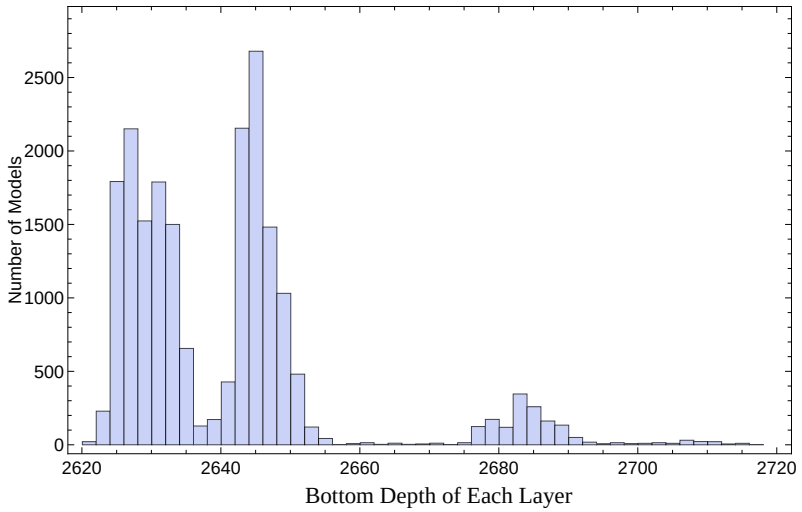


Fig. 3.24. The figure shows the total histogram of bottom depths of all chosen layers. As input variables, the number of layers = 3, the σ value = 1, the iteration=10,000, and the coefficient of step size=1, respectively.

Fig. 3.24 shows the total histogram of a bottom depth of each layer. From the result of the histogram, two portions of depths around 1630m and 1650m show a large number of models comparing with other portions, which means that many models chose these levels as a bottom depth of each layer. When upscaling well log data in an MCMC method, we need to carefully choose input variables because input variables such as the number of layers should be different depending on the distribution of well log data with depth.

The depth range also shows abundant hydrocarbon area from the evaluated data set regarding gas, oil, and water. From these results, the upscaling method by the MCMC algorithm can be freely applied to any depth areas we want to know. In order to more quantitatively analyze uncertainties in upscaling well log data, the specific relationships between a histogram and each input variable need to be examined with many experiments and statistical analyses.

3.8 Future Work

As seen in Fig. 3.23, the beginning portion of the well log data has a tendency to indicate physically more fluctuated properties comparing with other depth ranges, which means that these fine-scale depth ranges are very difficult to be seismically interpreted. In practice, detail information in well log data often hamper our correct decision for seismic interpretation because actual log data could contain unexpected errors or omit normal measurements for various reasons such as manual and instrumental errors. Therefore, those layers need to be upscaled as a more simple and pertinent blocked model as described in (b) of Fig. 3.23. The coarse-scale log data could contribute to making a better estimation of productive hydrocarbon areas.

In upscaling well log data which show rapidly varying velocity and density, some preliminary assumptions are proposed for simplified models. First, it is assumed that the media is horizontally stratified and the well is vertically measured. Therefore, we need to confirm wether an MCMC method can be applied to the cases of transversely anisotropic and azimuthal variations in layers or not, because Backus averaging is applied to only TI (transversely isotropic) models. In other words, it needs more research for cases of heterogeneity in orthorhombic symmetry.

Second, more advanced analyses for economic factors from MCMC results need to be considered. The quantification of uncertainties in upscaling will enhance the economic evaluation of reservoirs. To estimate the economic value, both of reservoir parameters (porosity, net pay thickness, water saturation, density, etc.) and economic parameters (oil price, production costs, interest rate, etc.) should be investigated and analyzed. In practice, economic factors such as IP (initial production rate) and NPV (net present value) have been statistically analyzed for years by oil majors.

Finally, the comparison of MCMC results with other seismic data (VSP or surface seismic data) could be added in future work. Because seismic data and well log

data use different source frequency, depth interval, and the number of layers, these differences could cause discrepancy between two results. Therefore, more specific studies such as travel-time corrections should be made continuously ahead.

CHAPTER IV

CONCLUSIONS

The main purposes of this thesis are efficiently upscaling well log data by geophysical inversion and statistically quantifying uncertainties in the upscaling. For the former goal, a Bayesian framework and a Markov Chain Monte Carlo (MCMC) method were used in the computational algorithm. For a theoretical background of upscaling, we need to understand that heterogeneities in rock properties can be considerably reduced by an averaging method such as Backus averaging. If a wavelength is large enough comparing with a layer thickness, seismic waves behave as if propagating in a transversely isotropic homogeneous medium, and this averaging makes it possible to replace a heterogeneous volume with a homogeneous volume. Most reservoirs have geologically heterogeneous structures from a few microns to several meters, which needs excessive time and efforts for many numerical simulations. Therefore, upscaling log data is a economical and effective process which is blocking thin layers into a thick layer keeping the same seismic properties of well log data.

This thesis is separated into three main parts to extend the previous work for upscaling to the quantification of uncertainties by an MCMC method.

In the first part, an MCMC algorithm based on a Bayesian framework was made for upscaling. Among many factors in the algorithm, the acceptance probability has the biggest influence on taking new models, and the value strongly depends on the ratio ($\frac{\triangle S^2}{2\sigma^2}$). In particular, a σ value is a key factor of controlling how many and how reliable models can be generated because the range of $\triangle S^2$ values is also determined according to the σ value. Conceptually speaking, artificial manipulation of a σ value means regulating how much 'noise' will be added in the given data.

In the second part, with respect to upscaling, the well log data of the North Sea were geologically analyzed with wave velocities (V_p, V_s), density, Poisson ratio, gamma ray, caliper log, porosity (gas, oil, water), etc., and the secondary properties such as V_p/V_s ratio and acoustic impedance showed close relationships with the well log data. As indicated in histograms of model parameters, an MCMC method yielded very reliable models in upscaling log data. In the actual experiment, it was proved that two very different seismograms in initial iteration changed into very close seismograms

after total iteration. In addition, the MCMC method yields many good models which make seismograms very similar to a reference seismogram, and the obtained models can be used for the quantification of uncertainties in upscaling. Furthermore, input variables (a σ value, iteration, the number of layers, and a coefficient of step size) can have an influence on quantity and quality of models.

In the last part, the total histogram of interface levels of chosen models showed comparatively larger number of models in approximate four portions nearby 1,650m, 1,695m, 1,710m, and 1,725m where indicate big variations in acoustic impedance. From the results, five layers can be selected for upscaling and the histogram of upscaling with five layers showed a more clear-cut distribution. Therefore, the histogram of interface levels can provide good information to confidently determine the number and location of blocked layers. Moreover, the MCMC upscaling method could be applied to various depth ranges containing abundant hydrocarbon areas.

REFERENCES

Anderson, D. L., 1961, Elastic wave propagation in layered anisotropic media: Geophys. Research, **66**, 2953–2964.

Backus, G. E., 1962, Long wave elastic anisotropy produced by horizontal layering: Journal of Geophysical Research, **67**, 4427–4440.

Bayuk, I. O., Ammerman, M., and Chesnokov, E. M., 2007, Upscaling of elastic properties of anisotropic sedimentary rocks: Geophysical Journal International, **172**, 842–860.

Cerveny, V., Molotkov, I. A., and Psencik, I., 1977, Ray method in seismology: Univ. Karlova.

Chi, X. G., and Han, D. H., 2006, Fluid property discrimination by AVO inversion: SEG/New Orleans 2006 Annual Meeting, pages 2052–2056.

Folstard, P., and Schoenberg, M., 1992, Low-frequency propagation through fine layering: 62nd SEG meeting (New Orleans), pages 1279–1281.

Gold, N., Shapiro, S. A., and Muller, T. M., 2000, An approach to upscaling for seismic waves in statistically isotropic heterogeneous elastic media: Geophysics, **65**, 1837–1850.

Helbig, K., 1958, Elastischen wellen in anisotropen medien: Geopys., **67**, 256–288.

Houck, R. T., 2002, Quantifying the uncertainty in an AVO interpretation: Geophysics, **67**, 117–125.

Hsu, K., Esmersoy, C., and Schoenber, M., 1988, Seismic velocities and anisotropy from high-resolution sonic logs: SEG Expanded Abstracts, **7**, 114.

Liner, C., and Fei, T., 2007, The Backus number: The Leading Edge, **26**, 420–426.

Mosegaard, K., and Tarantola, A., 1995, Monte Carlo sampling of solutions to inverse problems: Journal of Geophysical Research, **100**, 12,431–12,447.

Mosegaard, K., 1998, Resolution analysis of general inverse problems through inverse Monte Carlo sampling: Inverse Problems, **14**, 405– 426.

Mukerji, T., Avseth, P., Mavko, G., Takahashi, I., and Gonzalez, E. F., 2001, Statistical rock physics: Combining rock physics, information theory, and geostatistics to reduce uncertainty in seismic reservoir characterization: The Leading Edge, **MAR. 2001**, 313–319.

Sahu, S., 2000, Tutorial lectures on MCMC: University of Southampton, pages 1–46.

Schoenberg, M., and Muir, F., 1989, A calculus for finely layered media: Geophysics, **54**, 581–589.

Schoenberg, M., 1983, Reflection of elastic waves from periodically stratified media with interfacial slip: Geopysical Prospecting, **31**, 265–292.

Stovas, A., and Arntsen, B., 2006, Vertical propagation of low-frequency waves in finely layered media: Geophysics, **71**, T87–T94.

Tarantola, A., 1987, Inverse problem theory: Elsevier Science Publ. Co., Inc.

Thomson, W. T., 1950, Transmission of elastic waves through a stratified solid medium: Appl. Phys., **21**, 89–93.

Tiwary, D., 2007, Mathematical modeling and ultrasonic measurement of shale anisotropy and a comparison of upscaling methods from sonic to seismic: The University of Oklahoma.

Veire, H. H., Borgos, H. G., and Landrø, M., 2006, Stochastic inversion of pressure and saturation changes from time-lapse AVO data: Geophysics, **71**, C81–C92.

APPENDIX A

BAYESIAN FRAMEWORK FOR INVERSION

A fundamental purpose of Bayesian inference is to give us a quantitative solution from a posterior probability distribution of model parameters. The posterior information consists of the prior information preceding measured data and the likelihood quantifying misfits between measured data and calculated parameters. Prior to setting the prior distribution and the likelihood function, we should recognize that somewhat arbitrarily chosen prior and likelihood models sometimes could cause numerous unsolvable uncertainties in the posterior distribution. If inadequate prior models are input into an inverse problem, we cannot place much confidence on the posterior distribution. In fact, it is very crucial in an inverse problem to obtain reasonable solutions with the prior models because the posterior distribution is totally based on the prior information.

Bayes theorem is fundamentally expressed as a conditional probability, where this probability means a probability of an event occurring given evidence. In Bayesian estimation of a parameter, if a model parameter vector and an observed data vector is **m** and **d**, respectively, we could represent the posterior probability as following the Bayesian rule.

$$P(m|d) = \frac{P(d|m)P(m)}{P(d)} = \frac{likelihood \times prior}{evidence}, \tag{A.1}$$

where $P(m|d)$ is known as 'the posterior', and $P(d|m)$ is referred to 'the likelihood'.

To calculate the likelihood of the observed data, we should gather the prior information such as the distribution of parameters in an estimated model. Then, we can get the likelihood by multiplying the observed information by the prior information. The evidence term has a unit probability by normalizing over all values, and the posterior information could be also induced.

Sometimes, the Bayesian theorem seems to be a little controversial due to the strong dependency of the prior information and the difficulty of statistical assessment. Therefore, the prior information should be based on predictable and scientific information such as distribution of geological parameters.

When the observed data (d_{obs}) have Gaussian experimental uncertainties, the

likelihood function can be written as following.

$$P(d|m) = \frac{1}{(2\pi)^{n/2}|C_d|^{1/2}} \ exp[-\frac{1}{2}(g(m) - d_{obs})^T C_d^{-1}(g(m) - d_{obs})], \tag{A.2}$$

where *g(m)* is a calculated value from model parameters, and C_d is a data covariance matrix. If correlated errors are not included in the data, the covariance matrix C_d can be estimated as $C_d = \sigma^2\mathbf{I}$ in which $\mathbf{I}$ is an identity matrix, and the σ should be estimated as a standard deviation of the errors. If, of course, the error correlations exist without being neglected, the covariance matrix C_d must be applied in the inverse problem. The parameter set $\mathbf{m}$ means $\{m_1, m_2, m_3, \cdots\}$ in the model space and the measured data set represent as $\{d^1_{obs}, d^2_{obs}, d^3_{obs}, \cdots\}$ in the data space.

Concisely speaking, the likelihood function measures misfits between the observed data and the calculated values of model parameters. In case of upscaling log data, there are several parameters such as compressional and shear velocities, densities, and layer thicknesses in the function of a seismogram. Generally, the relationship between observed seismic data and calculated parameters can be described by a data model as following.

$$d = G(m) + e, \tag{A.3}$$

where $\mathbf{d}$ is the observed data and *G(m)* is the calculated parameters depending on the geometry of the measurement. In addition, the vector $\mathbf{e}$ means error terms based on noise or error. Theoretically, an inverse problem is to infer the parameters $\mathbf{m}$ with the measured data $\mathbf{d}$ and the error values. The posterior probability density (PPD) can be shown to be the integrated product of the prior times the likelihood according to the Bayesian rule.

$$P(d) = \int P(d|m)P(m)dm \tag{A.4}$$

This equation is to normalize marginal likelihood. Because the marginal likelihood is not a function of parameters, it is usually ignored for easier calculation.

A.1 Prior Information

The prior information is the prior belief of how some models are characterized in a specific system. This information gives a basis of available knowledge and general rock properties on reservoirs. Before applying the Bayesian inference, we need to spec-

ify the prior distributions for model parameters. Once model parameters are chosen, we may have the prior information from empirical data or other measurements. Based on this prior information, a range of parameters is determined with upper and lower boundaries. Each parameter, then, is usually uniformly distributed within this range. In practice, Gaussian approximation reasonably explains the relationship between the prior and the posterior information when we calculate the posterior probability.

A.2 Likelihood Function

The likelihood model contributes to links between real reservoir parameters and seismic data. That is, it measures misfits between observed data and calculated parameter values of chosen models. As mentioned in the equation A.2, the likelihood function is mainly expressed as a function of the measured and the predicted data. It is a very important procedure in Bayesian theory to derive the likelihood function from the relationship between available reservoir knowledge and the prior information consisting of regional geological information. One of the likelihood properties is that the total likelihood of a model can be expressed as a product of partial likelihoods: L(m)=L1(m)L2(m)$\cdots$, as one for each data type if data uncertainties are independent (Mosegaard and Tarantola, 1995). Furthermore, PDFs can be useful expressions to quantify stochastic model parameters from well log data.

A.3 Posterior Information

The posterior information combines the prior information with the likelihood function including new information from some observable data, and it shows a normal distribution if the prior and the likelihood have normal distributions, and the forward problem will be linear (Tarantola, 1987). Regardless of the aspects of the prior and the likelihood distribution, the solution form of the posterior is as following.

$$\sigma(m) = k\rho(m)L(m), \tag{A.5}$$

where **k** is a normalized constant, $\sigma(m)$ is a prior probability density, and *L(m)* is a likelihood function, respectively. The posterior distribution $\sigma(m)$ carries all available information on a model from measured data and prior information. From the posterior information, we can estimate better solutions and recognize a degree of

uncertainty in calculated parameters. In addition, useful information about model parameters satisfying the best-fit model can be obtained by the posterior distribution.

A.4 Application

In Bayesian framework, we can quantify uncertainties by using a simple way of deriving the posterior from the prior information and the likelihood function. For instance, stochastic reservoir models with fluctuating subsurface pressure or saturation can be quantified by a probability density function (PDF) of reservoir variables. Ultimately, the solution of this problem will provide quantification of uncertainties as well as direct estimations of the properties. In practice, PP reflection coefficient differences can be used to establish a likelihood model for linking reservoir variables with time-lapse seismic data. The methodology incorporates correlation between different variables of the model as well as spatial dependency for each of the variables. In addition, it will be possible to identify conflicts between prior model and measured seismic data and to evaluate the advantage of introducing new data (Veire et al., 2006). Chi and Han described that Bayesian AVO inversion method is a prestack inversion technique for inverting seismic elastic parameters such as P-wave velocity, S-wave velocity, and density (Chi and Han, 2006). The method is included in AVO processing and analysis, well-log editing and calibration, and prestack seismic inversion. In case of complex stochastic models, the differences between observed and calculated seismograms can cause complicated uncertainties, but also can be solved by Bayesian framework.

APPENDIX B

COMPARISON OF AN MCMC METHOD WITH A SIMPLE MC METHOD

An MCMC method has a computer-based algorithm relying on repetitive generation of random parameters. Although an MCMC method requires large amounts of random parameters, recent development of high-speed and huge-memory computer made it possible to generate numerous pseudo-random numbers promptly. A simple MC method is only to use randomly generated parameters independent of the state of previous parameters, whereas an MCMC method is to generate new parameters depending on a value of previous parameter. The theoretical approaches are very similar to each other, but the MCMC is more efficient in choosing more reasonable models because the next sample is always relevant to a present sample as much as a step size.

Example of a Polynomial Problem

B.1 A simple MC Method

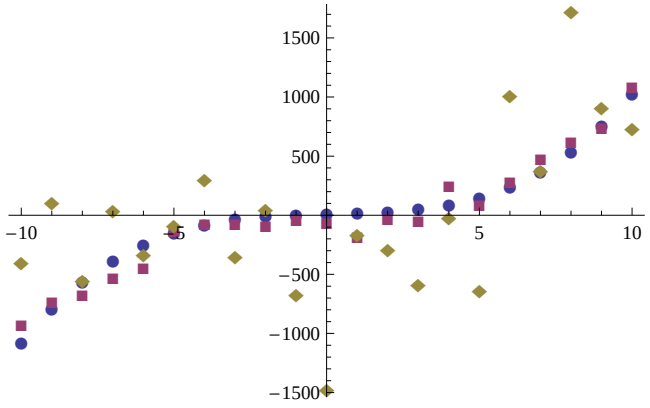


Fig. 6.1. The figure displays three blind data sets. The blue circle points show no error data, the red rectangular points are the data with a small error value added, and the yellow diamond points are the data with a large error value added, respectively.

As a simple inverse problem, we can consider an example of 4^{th} order polynomial

equation in order to understand a simple Monte Carlo method more easily.

$$f(x) = ax^4 + bx^3 + cx^2 + dx + e,$$

where coefficients *a,b,c,d*, and e are arbitrary numbers with the range of -10 to 10 and independent variable x varies with the same range. Assuming we have already known the observed data, our objective is to appropriately estimate a value of each unknown parameter *a,b,c,d*, and e satisfying the smallest error value. In addition, we intend to statistically analyze the results recognizing that the given data consist of three kinds of data sets; the exact, one with a small error value added, and one with a large error value added, respectively (see Fig. 6.1). These data sets were provided to the author as a blind test of the method.

For this purpose, a simple Monte Carlo method generating uniformly distributed random numbers was used with a fixed seed value. Strictly speaking, the random numbers are properly known as *pseudo-random* numbers; we shall refer to them simply as random numbers in computer programing. The word of 'random' for our purposes means that a sequence of such numbers should have the same probability of passing a statistical test for randomness as true random numbers would have (Mackeown 1997). In addition, it is very important concept for randomly generated numbers to be uniformly distributed. PDF of uniform distribution can be denoted as following.

$$p(x) = \frac{1}{(b-a)}, \qquad a \leq x \leq b$$

In the uniform distribution, mean and variance values are $\mu = (b-a)/2$ and $\sigma^2 = (b-a)^2/12$, respectively. With these uniformly distributed random numbers, we can calculate the error value which means the difference between the observed and the calculated data. With the condition satisfying the smallest error, we can obtain each fitting parameter in the polynomial equation.

The advantage of the simple MC method is that it gives a very easy and logically powerful algorithm, and what we should do is increase the number of iterations to get smaller error values and statistically analyzing the results to quantify uncertainties in parameters. Table 6.1 shows the result of each parameter satisfying the minimum error value in the given iteration. With increasing iteration, each parameter tends to converge into a specific value and to yield smaller error values.

In contrast, the method has two disadvantages. One is that it is an inefficient

Table 6.1. The table indicates the results of parameters when the calculated error value is smallest for the exact data during one million iteration. N^{th} means the iteration of the minimum error value (for example, what N^{th} is 5 means that the error value is smallest in 5^{th} iteration).

For Exact Data							
Iteration		10	100	1,000	10,000	100,000	1,000,000
parameter	a	-1.193660	-0.017936	-0.017936	0.045723	0.04572	0.00255
	b	1.284870	1.380540	1.380540	0.908644	0.90864	1.01162
	c	7.943110	8.315980	8.315980	-3.929290	-3.92929	-0.41253
	d	5.630920	-8.044650	-8.044650	2.836830	2.83683	-1.83218
	e	5.534570	-8.038070	-8.038070	1.495660	1.49566	9.40704
Error		20,049.6	1,697.0	1,697.0	326.8	326.8	193.7
N^{th}		5	62	62	6,497	6,497	682,246

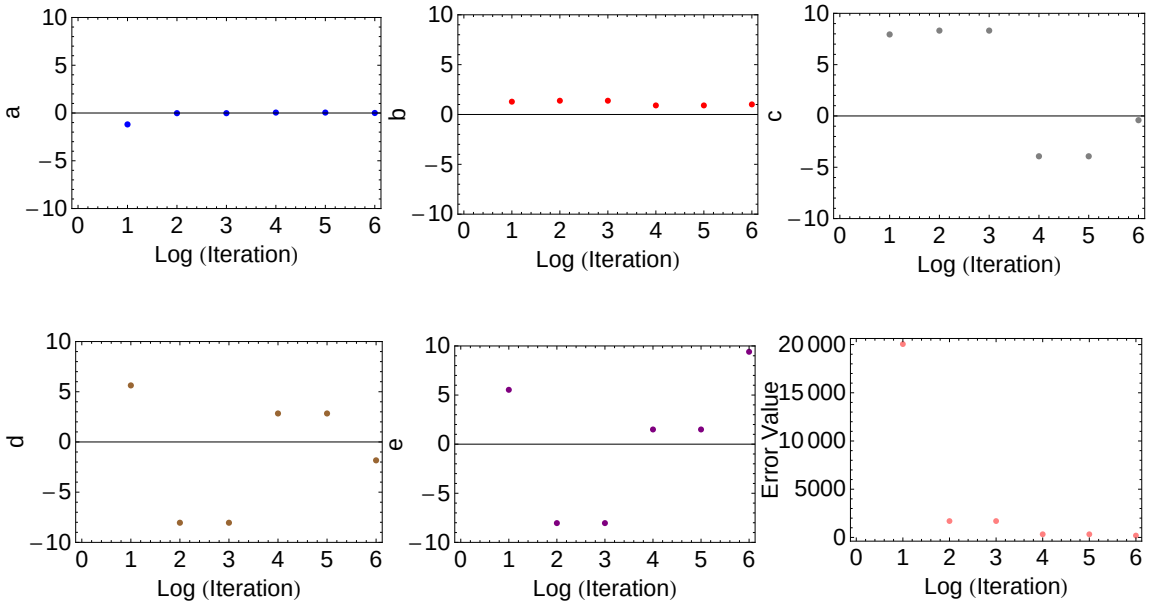


Fig. 6.2. The figures show the result of each parameter with increasing iteration for the exact data. The parameter 'a' and 'b' converge to 0 and 1, and the parameter 'c','d', and 'e' show fluctuating distribution. The error value goes down steeply with increasing iteration.

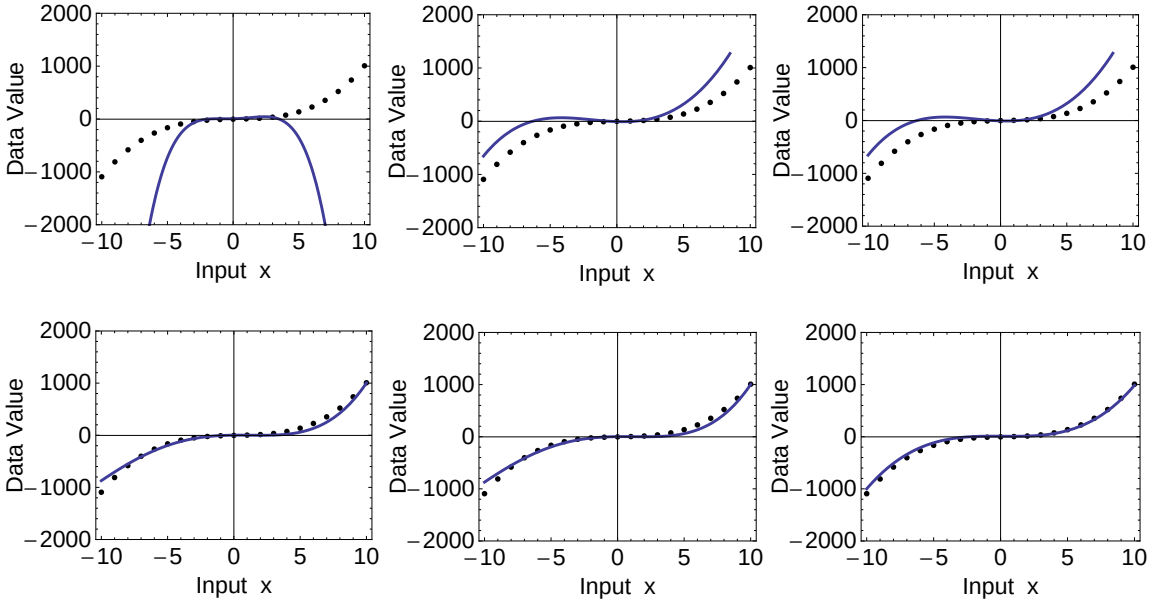


Fig. 6.3. For the exact data, the model prediction is approaching the given data (black points) as the iteration increases from 10 to 1,000,000. From the leftmost of upper figures, iterations are 10, 100, 1,000, 10,000, 100,000, and 1,000,000, respectively.

Table 6.2. The table indicates the results of parameters when the calculated error value is smallest for the small error data during one million iteration. N^{th} is the same as the one of table 6.1.

For Small Error Data							
Iteration		10	100	1,000	10,000	100,000	1,000,000
parameter	a	-1.193660	-0.017936	-0.017936	0.045723	0.04572	0.01728
	b	1.284870	1.380540	1.380540	0.908644	0.90864	1.08852
	c	7.943110	8.315980	8.315980	-3.929290	-3.92929	-1.13316
	d	5.630920	-8.044650	-8.044650	2.836830	2.83683	5.58022
	e	5.534570	-8.038070	-8.038070	1.495660	1.49566	-0.12338
Error		20,172.8	1,703.1	1,703.1	535.7	535.7	472.2
N^{th}		5	62	62	6,497	6,497	377,990

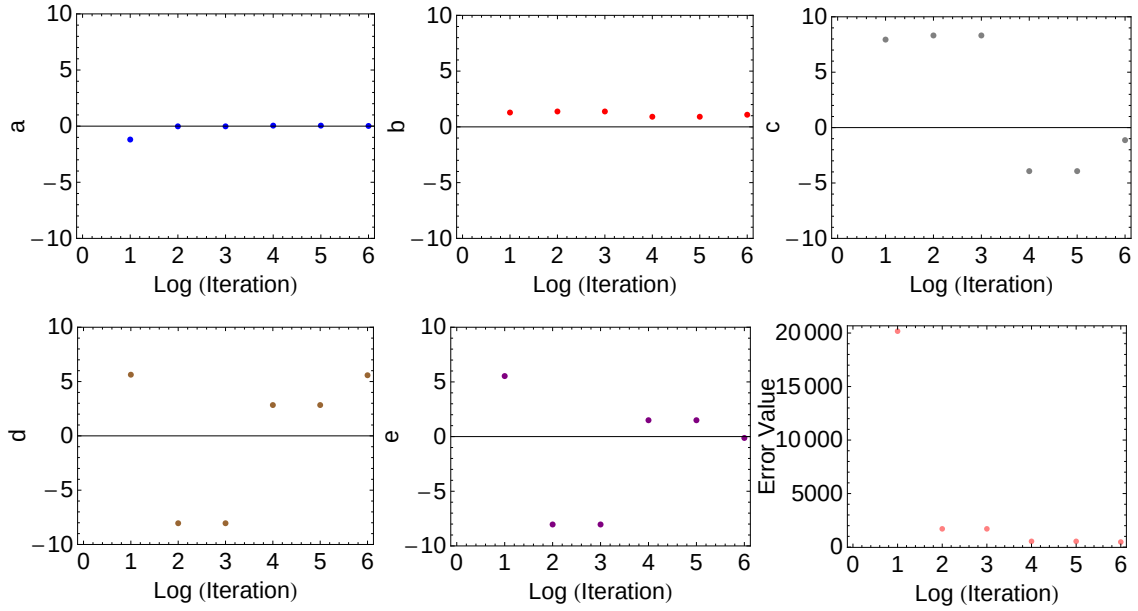


Fig. 6.4. The figures show the result of each parameter with increasing iteration for the small error data. The parameter 'a' and 'b' converge to 0 and 1, and the parameter 'c','d', and 'e' show fluctuating distribution. The error value goes down steeply with increasing iteration.

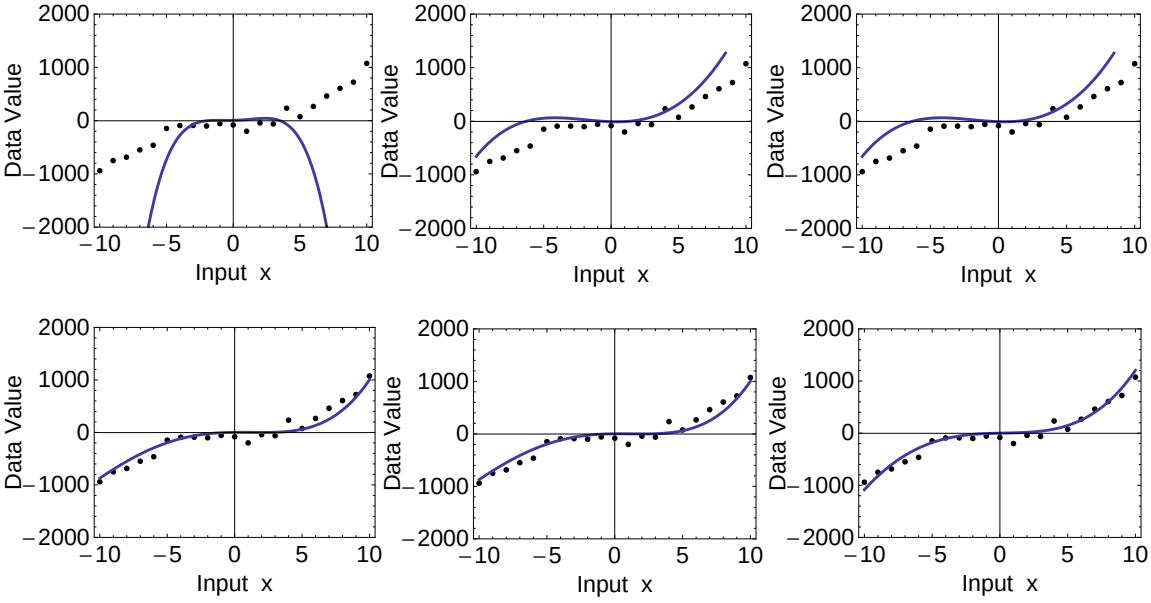


Fig. 6.5. For the small error data, the model prediction is approaching the given data (black points) as the iteration increases from 10 to 1,000,000. From the left-most of upper figures, iterations are 10, 100, 1,000, 10,000, 100,000, and 1,000,000, respectively.

Table 6.3. The table indicates the results of parameters when the calculated error value is smallest for the large error data during one million iteration. N^{th} is the same as the one of table 6.1.

For Large Error Data							
Iteration		10	100	1,000	10,000	100,000	1,000,000
parameter	a	-1.193660	-0.017936	-0.017936	-0.068301	-0.02611	-0.01556
	b	1.284870	1.380540	1.380540	0.361418	0.59752	0.75701
	c	7.943110	8.315980	8.315980	8.473920	4.78101	5.24229
	d	5.630920	-8.044650	-8.044650	8.191670	2.53070	8.35663
	e	5.534570	-8.038070	-8.038070	6.714570	-5.88085	-7.04301
Error		21,194.0	2,744.2	2,744.2	2,640.1	2,556.5	2,507.5
N^{th}		5	62	62	5,257	56,502	166,706

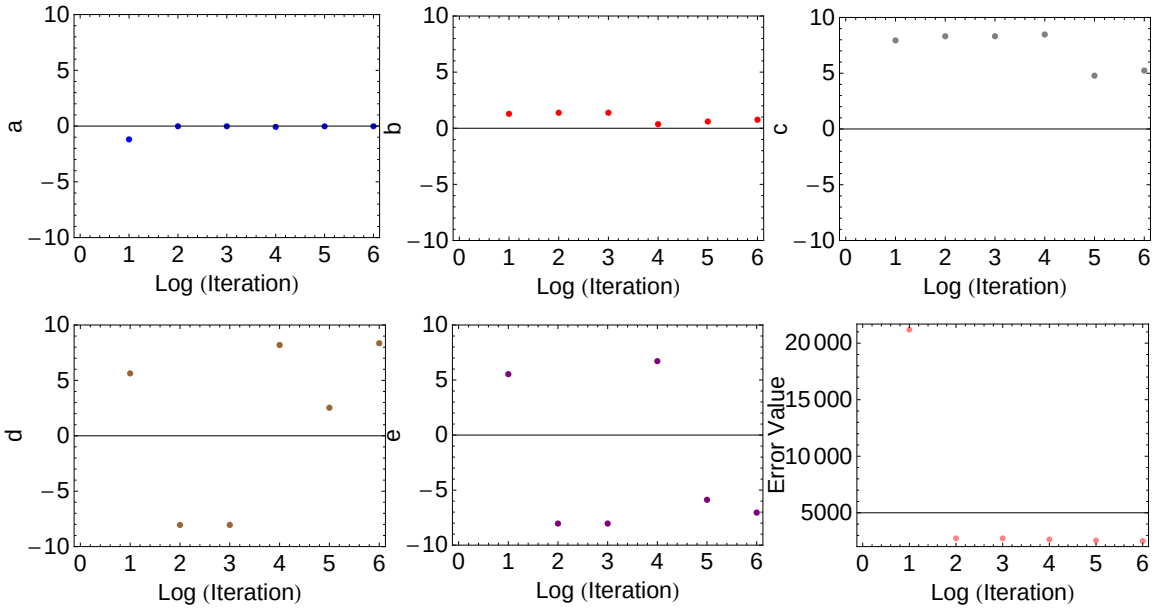


Fig. 6.6. The figures show the result of each parameter with increasing iteration for the large error data. The parameter 'a' and 'b' converge to 0 and 1, and the parameter 'c','d', and 'e' show fluctuating distribution. The error value goes down steeply with increasing iteration.

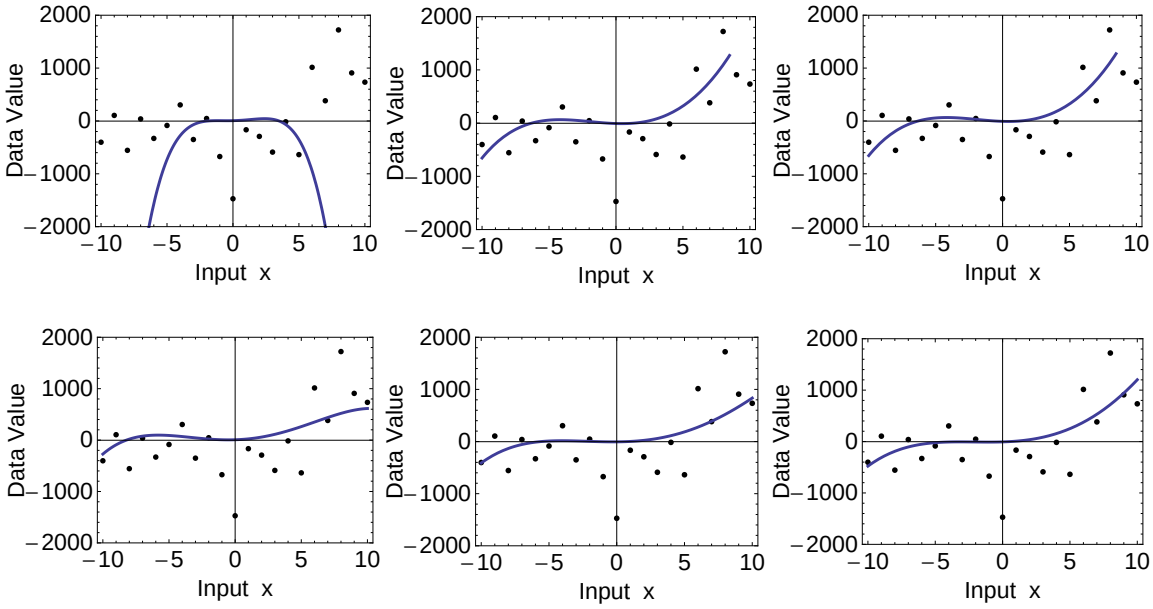


Fig. 6.7. For the large error data, the model prediction is approaching the given data (black points) as the iteration increases from 10 to 1,000,000. From the left-most of upper figures, iterations are 10, 100, 1,000, 10,000, 100,000, and 1,000,000, respectively.

method because randomly generated numbers are always independent of previous numbers. The other is that large iteration makes it take much longer time. In case of the polynomial equation, it is a very useful tool to estimate a simple model, but a more complex reservoir model with lots of geological properties such as densities, pore ratio, seismic velocity, layer thickness, etc. needs a more efficient way to obtain an optimum result.

The tables of 6.1 to 6.3 indicate quantitative results for each parameter and error value after one million iteration. From these results, error values for the exact data decrease more obviously than the ones of the large error data. Therefore, it is certain that recognizing how much 'noise' contain in the data is the most important problem to obtain reliable parameters.

B.2 A Main Algorithm of an MCMC Method

Prior to making a code, main steps of MCMC can be organized as below.

Step 1. The Problem

$$\begin{cases} d_{cal}(x) = g(m) = m_1x^4 + m_2x^3 + m_3x^2 + m_4x + m_5 \\ m_{ij} = [m_{1j}, m_{2j}, m_{3j}, m_{4j}, m_{5j}], j = [1, 2, 3, ..., n], n = iteration \\ d_{obs}(x) : \text{ Observed data to calculate the likelihood below,} \end{cases}$$

where x=[-10, -9,..., 9, 10], d_{cal} are calculated data values, and m_{ij} is parameter set.

Step 2. The Prior information

The parameters (m_{ij}) are uniformly distributed random numbers between -10 and 10.

Step 3. The rule to produce the next samples

$$m_{i(j+1)} = m_{ij} + \alpha \quad \begin{cases} i = [1, 2, 3, 4, 5] \\ j = [1, 2, 3, ..., n], \end{cases}$$

where α (known as step size in MCMC) is uniformly distributed random number between -1 and 1, and $m_{i(j+1)}$ is the next model of m_{ij}.

Step 4. The criterion to decide which model is better or worse

$$L(m_{ij}) = k\ exp[-\frac{1}{2}\frac{S^2(m_{ij})}{\sigma^2}] \tag{B.1}$$

$$S(m_{ij}) = [\sum_{x=-10}^{10} (d_{cal}(x) - d_{obs}(x))^2]^{\frac{1}{2}}$$

$$m_{i(j+1)} = \begin{cases} \text{Better Model} & \text{if } L(m_{i(j+1)}) \geq L(m_{ij}) \\ \text{Worse Model} & \text{if } L(m_{i(j+1)}) < L(m_{ij}), \end{cases}$$

where σ^2 means 'noise' variance, $L(m_{ij})$ is a 'Likelihood' from model parameter m_{ij}, and $S(m_{ij})$ is error term from measured data and calculated data.

Step 5. Generate the posterior models with the comparison between computed and observed data using the Metropolis rule.

For the acceptance probability(P_{accept}),

$$P_{accept} = \begin{cases} 1 & \text{if} \quad L(m_{i(j+1)}) \geq L(m_{ij}) \; ----(1) \\ \exp(-\frac{\triangle S^2}{2\sigma^2}) & \text{if} \quad L(m_{i(j+1)}) < L(m_{ij}) \; ----(2), \end{cases} \tag{B.2}$$

where $\triangle S^2 = S^2(m_{i(j+1)}) - S^2(m_{ij})$.

$$m_{i(j+1)} = \begin{cases} m_{1j}, m_{2j}, m_{3j}, m_{4j}, m_{5j} & \text{in case of (1)} \\ m_{1(j+1)}, m_{2(j+1)}, m_{3(j+1)}, m_{4(j+1)}, m_{5(j+1)} & \text{in case of (2),} \end{cases}$$

where j=[1,2,3,...,n], n=iteration.

$L(m_{ij})$ and $L(m_{i(j+1)})$ are calculated from m_{ij} with $[m_{1j}, m_{2j}, m_{3j}, m_{4j}, m_{5j}]$ and $m_{i(j+1)}$ with $[m_{1(j+1)}, m_{2(j+1)}, m_{3(j+1)}, m_{4(j+1)}, m_{5(j+1)}]$, respectively.

B.3 Results of MCMC

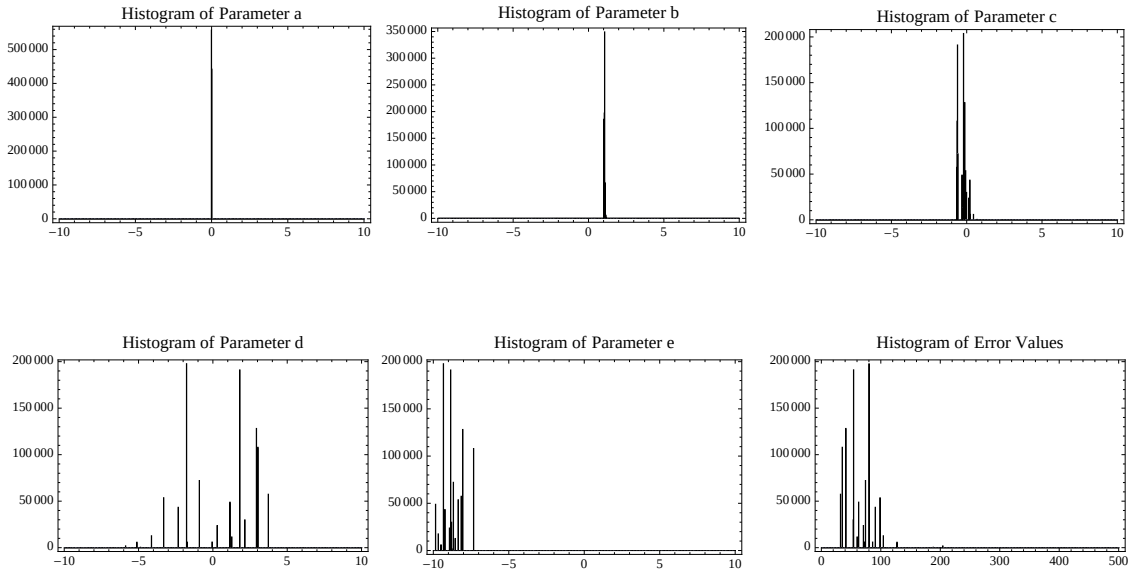


Fig. 6.8. The figures indicate the histograms of the parameters and the error value after one million iteration. The MCMC method could provide the histograms of parameters, which makes uncertainties in the results quantified.

Histogram is a representative result of an MCMC method because we could understand how widely the results are distributed and how much models are chosen with specific values. Fig. 6.8 indicates the histograms of parameter and error value.

In case of the parameter ‘a’, most results are concentrated on the value zero, which means that most generated models have an almost same value independent of other parameter values. From the given polynomial equation, we could presumably expect that the parameter ‘a’ is the most influential coefficient to considerably effect on the result of the polynomial equation. The histogram of the parameter ‘b’ indicates a similar tendency except that most results are distributed on the different value 1. The parameter ‘c’, ‘d’, and ‘e’ shows wider distributions comparing with the parameter ‘a’ or ‘b’. These histograms could give us a good confidence in estimating parameter values to some extent. Therefore, an MCMC method gives us good estimation when analyzing significant uncertainties in stochastic models with a statistical analysis such as histograms of model parameters. In addition, this method is very useful when it is infeasible or impossible to obtain exact solutions with a deterministic system.

VITA

Kyubum Hwang received a B.S. degree in Civil Engineering from Yonsei University (1998). He had worked as a supervisor in the fields of design and construction of LNG underground Tank for eight years. He joined the research group of Dr. Richard Gibson (Aug 2007) and received an M.S degree in Geophysics (May 2009). He is now working in G&G (Geology & Geophysics) Team of Korea Gas Corporation, 215 Jeongja-dong, Bundang-gu, Seongnam-si, Gyeonggi-do, South Korea, 463-754. His email address is kyubumok@gmail.com

The typist for this thesis was Kyubum Hwang.